

22 April 1982

Zeal on secrecy could backfire

Admiral Bobby Inman of the CIA knows the danger that his crusade to alert US researchers to Soviet espionage could make academics abreast. But what else is the Administration looking for?

The Reagan Administration may be heading for a confrontation with US universities reminiscent of the Vietnam era. The issue is the tradeoff between national security and academic freedom in the research universities and the charge by Administration officials that, if something is not done, US campuses will become the targets of intensified Soviet intelligence efforts to gain advanced knowledge in genetic engineering, very high speed integrated circuits, lasers and the like. So far, the dialogue between the Administration and several spokesmen for university science has been restrained, even cordial. But there is already some hardening of positions on both sides, and some Administration ire that the scientists are not cooperating, as well as some campus resentment of the calls for self-policing.

Some of this resentment surfaced on 29 March when two sub-committees of the House Science and Technology Committee of the US Congress held a joint hearing to gather the Administration's view of how the free flow of research threatens national security. The star witness was Admiral Bobby R. Inman, the deputy director of the Central Intelligence Agency (CIA). Inman was formerly director of the National Security Agency (NSA), the government's secret code agency, where he provoked a dialogue with university mathematicians on the dangers of publishing advances in cryptography — and even got some to submit papers voluntarily for government censorship. In January of this year, Inman caused a stir when he ventured to the annual meeting of the American Association for the Advancement of Science and suggested that unless scientists themselves devised mechanisms to prevent leaks to the Soviet Union, the government might require them to submit their research to censorship before publication.

At the House hearing in March, Inman elaborated his view of Soviet technical espionage in the United States, US counter-espionage and the campuses. At present, he said, approximately 70 per cent of Soviet technical intelligence comes through purchases of hardware — some legal, some illegal. Moreover, the Soviet Union does "a very thorough vacuum cleaning of anything in the public sector". Finally, there is a smaller effort, which has received special emphasis since the late 1970s, aimed at acquisition of technologies emerging from universities and research centres. US counter-espionage is aiming to cut off as much of the hardware acquisition as possible. So, Inman predicts, US campuses will become the focus of a much intensive espionage. In 6 months, a year or 18 months, he warned at the hearing, it will become clear to everyone, including the Administration, how much the Soviet Union is benefitting from the "outflow" of information from the campuses on some sensitive topics. Pressure will then grow in the Administration to crack down on universities — unless by then, they themselves have controls in place.

Inman's role in this crusade is odd. In March, he repeated that he was not representing the CIA or the Administration and that his testimony had not been coordinated with the other Administration witnesses. He described himself as a self-appointed "goat" or "gadfly" to get people to think about the problem and to suggest "innovating ideas" for cutting down the "outflow" from their laboratories to the Soviet Union. The trouble with his case, so far, is that few university scientists seem convinced there is a large outflow or "haemorrhaging" of information about very high speed integrated circuits, robotics or anything else. Inman admits his case rests on evidence contained in a classified report that cannot be released without compromising intelligence sources. But representative Albert Gore Jr (Democrat, Tennessee), chairman of the committee's sub-committee on

investigations and oversight, who has seen the classified report, ventured at the hearing that the full evidence is not all that convincing. Inman himself conceded that "In an open society, there is a tendency to automatically (*sic*) take the view that one cannot debate the substance of issues unless you see the evidence." He and the Administration should understand that the discussion of what the universities should do cannot begin until the evidence is on the table.

The Administration's new-found zeal for secrecy about academic research has already had some tangible consequences. Some months ago, on the eve of a visit planned for a Soviet robotics expert to the United States, the Administration asked Stanford University and some other places to close their electronics laboratories to him. The universities replied that they could not police their visitors, whereupon the would-be visitor was not issued with a visa. The truth is that US universities cannot easily act as espionage policemen, especially when they are trying to forge links with industry to give US industry an "outflow" of

Visitor: Here's \$250,000 to spend. No strings. Just let me be the first to know.

Scientist: Hey! Don't you have relatives in Eastern Europe? How do I know you're not a spy?

Visitor: That's none of your business. Anyway, I'd been thinking of making a deal with that competitor of yours in Berkeley. How do I get out of here?

research results it badly needs to stay competitive with Western Europe and Japan. Does Inman really want US scientists to be suspicious that every foreign visitor or industrial benefactor may be a spy? (See box.) Can it be wise to breed such suspicions when university scientists are making their first fragile attempts to reach outside their laboratories?

In reality, however, the issues the Administration has raised with the campuses offer much middle ground and room for discussion. University scientists already live with a spectrum of controls — from peer review to federal regulations — that appear not to inhibit academic freedom. More and more people are learning about the new shadings of controls required when dealing closely with industry — the distinction, for example, between teaching only basic science in the classroom and keeping possible patentable discoveries off limits for a certain period. Moreover, the Administration is engaged on several reviews that could have enormous impact on university research. The Commerce Department is reviewing the International Traffic in Arms Regulations (ITAR), passed in the Cold War era of 1955 (and which is probably unconstitutional), that allows the government to prevent to the export, and possibly publication, of "technical data" — whatever that is. At the hearing in March, in chilling testimony, Assistant Secretary for Trade Administration in the department, Lawrence J. Brady, said that in the past, it has not been fully understood that ITAR applies to all segments of US society, including the universities. George Millburn, an under-secretary at the Department of Defense, testified that a "struggle" is going on in the department to define which technologies are militarily important and subject to review and control. Yet the only academic involvement in these discussions is some work by the Defense Science Board and a study by the National Research Council of the National Academy of Sciences



(see page 695).

While there may be room for discussion, there are also the makings of a fight. Many academics or their champions, such as Gore, see in the Administration's words a threat to fundamental inquiry. Gore told Inman "we don't want to even (*sic*) take the first step along the road that has made Soviet science so pitiful". Inman, on his side, complained that the press and some academics misrepresented his statements to mean that he wanted to "throw a net over the public". Yet some universities have already taken avoiding action. At Stanford University the Faculty Senate has approved a resolution urging the university to resist interference with free scientific communication.

Academics are highly sensitive to any threats to their independence and tend to cry foul when anyone suggests they submit to any authority whatever — be it university administrators trying to enforce government accounting rules or suggested self-censorship of militarily important research. Moreover, President Reagan's military build-up and talk of nuclear war is stirring antagonism at many US campuses, as the recent wave of teach-ins and meetings attests. It would be all too easy for these issues to be rolled into one ball of wax in many academics' minds: burdensome federal regulations, arms build-up, nuclear war and government controls to prevent espionage. The resulting feeling would be that the Administration is against the universities and does not understand them — a feeling reminiscent of some campus attitudes towards former Presidents Johnson and Nixon. Given its penchant for loose, inflammatory rhetoric on other matters, the Administration may find it too easy to characterize campus resistance to controls as unpatriotic. The push from the right would then shove the campuses left. If the Administration keeps to Inman's timetable, the universities will have to step briskly — and calmly and rationally — to avoid an open conflict.

Big Brother's law

The British Government seems indifferent to its electors' needs for protection from data banks.

The British Government has simply failed to understand why there is public anxiety that computerized records may be a threat to people's privacy and, thus, liberty. That is the most charitable reading of the white paper on the subject, now at last published.* A more cynical, but probably more accurate, interpretation is that the government is squarely out of sympathy with the fear that computerized data banks may infringe the rights of individuals for which reason it will continue to drag its feet on legislation, doing everything it can to ensure that the eventual law is the bare minimum necessary to comply with the Council of Europe Convention signed a year ago.

The grudging character of the government's intentions can be told from the only statement in the white paper of the reasons why legislation is necessary. This mean document correctly records that the Council of Europe convention will require its signatories not to transmit personal data files to countries without satisfactory legislation, notes that this restriction could be a threat to international companies operating in Britain as well as to British computer bureaux and therefore concludes that "in order to conform with international standards of privacy protection and to avoid possible barriers to trade", the government will introduce legislation that will "enable the United Kingdom to ratify the convention".

Shabbily, nowhere is there a flicker of acknowledgement that the protection of people's privacy is in itself desirable, a public good of the kind which governments are elected to cultivate. Indeed, the government's sense that it owes something to its electors is expressed only in the repeated promise that the cost of administering the promised legislation will be kept to a minimum, and will be no charge on public funds. Even the salaries of the proposed registrar of personal data banks and his staff will be recovered from the fees that must be paid by their operators. Individuals seeking to verify the accuracy of personal data on

specific computer files will have to pay a fee "based on the principle that the costs for the demands are fully recovered". Those who, having paid their fees, think they have found their privacy to have been illegally infringed will have to mount prosecutions off their own bats, and at their own expense.

The government that happens historically to owe its existence to the *Magna Carta*, one of the earliest public declarations that tyranny is unacceptable to ordinary people, appears to be willing that ordinary people should take their chances in the courts if they have reason to believe that some injustice has been done by a data bank or those who operate it. Fortunately — this at least must be the hope — even the government's supporters in the House of Commons will be so affronted by what is now proposed that the unseemly legislation the government plans belatedly to introduce (not this year, perhaps not even next year) will be constructively and radically amended.

This is how libertarians in the House of Commons and elsewhere should argue. While computerized data banks are in principle no different from systems of records manually kept, they are potentially so much more efficient that their common use is tantamount to a qualitative change. So much has been acknowledged in the past decade by two government committees set up specifically to consider questions of privacy — the Younger committee (which reported in 1972) and the Lindop committee (reporting in 1978). False data, inadvertently or even maliciously erroneous, can be damaging to those concerned. But even accurate data can be damaging, especially in the hands of those for whom they were not originally compiled. While it may be proper that information about people's dealings with tradespeople should be available to others in the same line of business, information such as this may be unwarrantably damaging in the hands of others, potential employers for example. Similarly, information about a person's medical history collected by one government department in the course of its legitimate business may be damaging when transferred to another.

So how should this transfer of data between different kinds of users be regulated? In its declaration of the principles on which the new legislation will be based, the British Government merely quotes the principles on which the European convention has been drawn (the chief of which is that a person should have access to any files about him, and the power to correct them if inaccurate or to erase them if illegally compiled). What needs to be decided, now and not later, is the basis on which the transfer of data may be allowed from one potential user to another. It is simply not good enough to say that "it is expected that most applicants will be registered without question, but the Registrar will have power to make enquiries, to inspect data files and to require modifications to a system. In extreme cases, he may need to refuse registration. . .". What the government has so far failed to explain is how the registrar will decide which applications for registration are acceptable "without question".

The issue is especially important where the transfer of information between government departments is concerned. The European convention permits the exclusion from regulation of data banks "protecting state security, public safety, the monetary interests of the state or the suppression of criminal offences". This form of words is a sufficient licence for Big Brother. On the face of things, it will for example allow people's income-tax returns to be compared with their income from holdings of government securities or their applications for driving licenses to be checked against the records of the National Health Service, while the police will presumably be allowed to gather information from where they can in the cause of crime prevention. Are all these applications of data banks to be blithely sanctioned? By its repeated references to the European convention as the guiding principles on which the registrar will work, the British Government seems to intend just that. Parliament should insist that the European convention is not the final touchstone of what is permissible, but merely the starting point for what is bound to be a long and acrimonious argument.

*Data Protection, Cmnd 8539, 24pp. (HMSO, London, 1982) £2.30.

Oppenheimer case boils up again

Files show FBI bugged pre-trial conversations

Washington

A forthcoming history of science journal reveals new information on the celebrated Oppenheimer case — the most traumatic incident in post-war US science — that suggests that the government stacked the case against Oppenheimer in proceedings that caused the revocation of his security clearance and his personal disgrace.

Physicist J. Robert Oppenheimer led the Los Alamos atomic bomb project during the Second World War. Afterwards, he was the unquestioned leader among US physicists seeking to shape the post-war world. Oppenheimer was influential in the debate over civilian control of atomic energy and, later, chairman of the General Advisory Committee of the Atomic Energy Commission (AEC). Oppenheimer was then without doubt the government's most influential science adviser.

But doubts about his possible pre-war Communist associations, and his apparent evasions in relation to them, led AEC to conduct an inquiry into his suitability to hold security clearance, and to a hearing before a specially appointed triumvirate of AEC's Personnel Security Board in 1954. The hearing, including the testimony of his one time colleague Edward Teller, destroyed him.

An article by Barton J. Bernstein, Professor of History at Stanford University, in *Historical Studies in the Physical Sciences**, to be published later this spring, alleges that despite President Eisenhower's stated wish that the government should conduct the inquiry in a dignified manner, commensurate with Oppenheimer's stature, the Federal Bureau of Investigation (FBI) nonetheless bugged Oppenheimer's conversations with his attorneys before the inquiry and passed the transcripts on to prosecutor Roger Robb. AEC chairman Lewis Strauss personally thanked the FBI for their help in supplying the transcripts. "It would be an uneven contest", even before the hearing, writes Bernstein, "just as if one army had stolen its enemy's secret plans".

Moreover, Bernstein says, while members of the board and the prosecution had access to secret FBI files on Oppenheimer (some of which were unopened at the time that the government had given Oppenheimer earlier security clearances), these were never shown to

Oppenheimer or his attorneys on the grounds that security clearance could not be obtained in time.

These and other revelations are based on documents declassified by the FBI, the Department of Energy (successor to AEC) and the Eisenhower Library at Bernstein's request. But Bernstein believes there is still more to learn about the case, as "more than half" of the FBI material on Oppenheimer remains classified.

The chairman of the three-member panel that finally judged Oppenheimer a security risk was Gordon Gray, former president of the University of North Carolina and former Secretary of the Army. Gray's papers, unearthed by Bernstein, show that Gray considered the other two members of the board to be completely prejudiced against Oppenheimer before the hearing even started. The three sat down together beforehand, to go through the files. Gray wrote of the other two, "It seemed to me from the beginning that Mr Morgan and Dr Evans both had a strong hunch that Dr Oppenheimer's clearance should not be reinstated . . . During the weeks spent reading through the files this notion became much clearer." Evans made anti-Semitic comments, saying that he found "that nearly all subversives were Jews". "At the time I was concerned at this note of clear prejudice", Gray wrote. But he did not tell AEC that the board's members were not impartial as AEC rules required.

Bernstein's new documents also reveal that Edward Teller, whose testimony was crucial, had told the FBI some years earlier of his doubts about Oppenheimer, and that Teller's statements helped shape the charges. Teller, however, has said that he was undecided about Oppenheimer until

the night before he testified, when Robb showed him Oppenheimer's security record.

But in 1952, according to Bernstein's article, Teller told the FBI that he would do almost anything to separate Oppenheimer from the General Advisory Committee because Oppenheimer had cleverly shifted from one argument to another on the committee and used his considerable influence over other scientists to delay development of the hydrogen bomb. In 1952, Teller had not said that Oppenheimer was disloyal, only that he was dangerous because he gave bad advice. To Teller in 1952, Bernstein writes Oppenheimer "was devious but not disloyal". He had persuaded Hans Bethe to refuse to work on the H-bomb and had swayed Henry D. Smyth, an AEC commissioner, to oppose it.

Bernstein believes that his new information also provides an even more "alloyed" portrait of Oppenheimer himself than has emerged in previous writing about the case, which has tended to portray Oppenheimer as a martyr. The documents show, Bernstein writes, evidence of "a string of lies and evasions" regarding his past, and that Oppenheimer had not always been candid with the authorities or his associates. Commenting on the strain that Oppenheimer visibly underwent at the hearing, Bernstein writes:

Oppenheimer was struggling to preserve his reputation . . . He lost. And he may have been under an additional strain that his defenders did not appreciate, for he may also have been hiding his former membership in the Communist party. Chevalier claimed that Oppenheimer had been a member, and Robb reached the same conclusion.

However, except for wondering what is in the yet unopened files, Bernstein says there

Merck backs UK neuroscience

Dr Leslie Iversen, director of the Medical Research Council's neurochemical pharmacology unit at Cambridge for the past eleven years, has been appointed director of a new laboratory being established in Britain by Merck, Sharp and Dohme. The new laboratory, which will be a going concern on a green-field site near Harlow, Essex, by 1984, will be exclusively concerned with basic drug-oriented research in the neurosciences and will eventually employ 200 people, a third of them with research degrees.

The British scientific community is delighted at this development, chiefly because of the jobs it will provide for researchers but also because it is a token of continuing American interest in the United Kingdom as a base for major research laboratories.

Dr Iversen said earlier this week that he was looking forward to the job, but emphasized that while the new research

centre would concentrate largely on basic neuroscience, it would be an integral part of the company's research effort and thus different from the Roche research centres in Switzerland and the United States. The hope, he said, was that basic neuroscience research would suggest novel classes of therapeutic drugs.

The only cloud on the horizon is the future of the Medical Research Council's unit, which is to be decided in June this year and which has been for some years one of the brightest feathers in the council's cap. Dr Iversen says that he is anxious that the unit should survive his departure on 1 March next year, for which reason he intends to avoid recruiting any of his present staff until the council has made up its mind. Apart from the unit's basic research, it also houses the United Kingdom "brain bank" from which authenticated samples of brain tissue are distributed to qualified investigators. ●

**Historical Studies in the Physical Sciences*, Office for the History of Science and Technology, 470 Stevens Hall, University of California, Berkeley, California 94720, USA.

is no direct evidence in the new material that Oppenheimer had indeed been a Communist.

Deborah Shapley

Soviet radio-catastrophe

Another theory

Washington

What *did* happen in the southern Ural mountains in 1957–58? Did buried radioactive waste explode? Or was the landscape blighted and thousands of people hurt by a strange weather pattern that rained fallout from Soviet atomic bomb tests in the north?

The disaster has been a mystery since it surfaced in the West in a Copenhagen newspaper report in 1958 that implied that a Soviet weapons test had gone awry, and speculation grew into controversy in 1976 when Zhores Medvedev, a prominent Soviet dissident scientist living in London, alleged there had been a catastrophic, volcano-like explosion from buried nuclear wastes in a secret area near the city of Kyshtym.

Now a new theory has been advanced by the Los Alamos Scientific Laboratory, that links the disaster, among other things, with acid rain. The public summary of a classified report concludes that there was no single catastrophic explosion of buried wastes, and certainly not an explosion involving nuclear criticality. Instead, widespread damage to the land and population was caused by "a series of relatively minor incidents . . . severely compounded by a history of sloppy practices" at a plutonium production plant.

The new report confirms the allegations of Medvedev and Soviet scientist Lev Tumerman, who drove through the region in 1960, in their horrid accounts of burned people and devastated land. It recounts that there were "death squads" of radioactively contaminated prisoners assigned to work in the area, essentially until they dropped dead.

The site of the alleged disaster was a plutonium production plant built for the Soviet weapons programme east of the city of Kyshtym on the site of a munitions plant that operated under the Tsars. The plutonium plant was apparently built in the late 1940s, after the Smyth and Groves reports had been published explaining how the United States had built its atomic weapons during the war. The report indicates that the Soviet Kyshtym plant was modelled on the US plutonium plant at Hanford, Washington, and that, the Soviet Union had detailed information about Hanford practices and procedures. The evidence for this, however, is not given.

So, the report says, a city was built having "thousands" of workmen and engineers. A graphite-moderated, uranium-fuelled reactor was built, called "Unit O". Cooling water from Unit O was pumped into a large artificial lake dug on the grounds. Later, the cooling water from

Rorvik disowned

New York

Dr Derek Bromhall, formerly of the University of Oxford, has won his suit against journalist David Rorvik and the publisher Lippincott and Company. On the third day of the trial this month (see *Nature* 1 April, p.383) and before a single scientist had taken the witness stand, the defendants settled out of court. The settlement is important not only because it amounts to a public disavowal by the publishers of Rorvik's book, published in 1978 under the title *In His Image: the Cloning of a Human*, but also because it is a precedent for other complaints of the unauthorized use of scientific data in popular works.

In an apology to Dr Bromhall which accompanied the settlement (said to amount to some hundreds of thousands of dollars) Martin H. Lippincott, Chief Executive Officer of the publishers, conceded that "this company now believe the story to be untrue" and acknowledged that "Dr Bromhall did not consent to the inclusion of his name or his research techniques in the book and also that Dr Bromhall never engaged in, or attempted to engage, or advocated the cloning of a human being. We regret any embarrassment, humiliation, or other injury to him that the reference to Dr Bromhall in the book might have caused".

Lippincott and Company had promoted the book, about an unnamed American millionaire financing the birth of his clone, as non-fiction.

Judge John Fullam of the United States District Court, who had been hearing Dr Bromhall's suit last year, called the book "a fraud and a hoax". Apparently Lippincott began withholding royalty cheques from Mr Rorvik as soon as the initial suit began.

As Dr Bromhall is no longer an academic scientist, the direct damage to him from the book where his cloning techniques were quoted would have been difficult to prove, according to observers. Only the threat of punitive damages forced the hasty end to the trial.

Michael Stein

at least two reactors built there was also pumped into the lake. Fuel-rod cladding failures were probably frequent, the report says, so the rods leaked material. By 1953, the lake had become dangerously radioactive.

The lake fed into the Techa River, a slow-moving stream along which lived a primitive Bashkir people. Soon, the Techa, too, became "a chronic, waterborne source of radioactive contamination" long before any of the dates of alleged disasters. To separate plutonium from the irradiated fuel elements would have needed nitric acid. The off-gas belched yellow smoke out from a high smokestack, 24 hours a day for

years on end. Because of the high humidity of the region, the Los Alamos theory goes, the nitrogen oxides in the gas would have produced acid rain and acid-laden snow which could have caused the "dead" landscape seen by Tumerman.

The Soviet Union was in such a hurry to get as much bomb-grade material as possible that the irradiated fuel elements were not allowed to cool, so the gases from the plant would have included radioactive iodine as well. The authors of the report say that the impact of the consequent acid rain would have been appalling. A third source of contamination was the burial of radioactive wastes from the late 1940s and the 1950s. The wastes were not put in steel tanks as at Hanford but radioactive liquid waste was simply dumped into a dry lakebed 5 or 6 kilometres south-east of the plutonium complex.

People in the Techa River region began having radiation sickness, crops were killed by acid rain and the stackgases. So, the report says, the authorities must have evacuated the Bashkirs and burned their homes. This would account for the burned buildings, and for the disappearance of local villages — signs Medvedev interpreted as indications of a nuclear disaster.

The Los Alamos report, like others, confirms that the landscape of the region was changed mightily in the late 1950s. The report includes "before" and "after" maps showing that new lakes were created, drainage canals dug to circumvent other lakes, and so on. All this, it seems, was necessary to stem the contamination of the Techa.

If there were explosions, the report concludes, they would have been chemical, not nuclear. Likely causes cited include ammonium nitrate and hexone, a flammable solvent often used in waste separation.

The report is the latest in a chain of Western "explanations" for the disaster. Curiously, one of the report's two authors, Danny Stillman a Los Alamos staff member, was a co-author of an earlier Los Alamos report attributing the disaster to fallout from "dirty" atomic weapons tests to the north at Novaya Zemlya, and a meteorological fluke that dumped the fallout southwards. Stratton was not available to discuss why the theory has changed: nor was the other author, Diane M. Soran. However, Harold Agnew, a former Los Alamos director who is president of General Atomic Corporation, and who co-authored the Novaya Zemlya theory, says the new theory is an improvement.

The proponents of the buried nuclear waste explosion theory are not convinced. Medvedev says that the report is simply wrong. The people who lived in the area were not simple Bashkir tribesmen as claimed in the report — the Bashkir Republic is somewhere else: the villages that disappeared after the disaster all have Russian names, and therefore had Russian inhabitants.

Moreover, Medvedev says his notion all

along was that the nuclear waste had heated and exploded thermally, or perhaps chemically. Especially during the first two years, nuclear wastes can heat up and possibly explode, unless they are carefully cooled. Medvedev notes that his theory, and a study published by the US Oak Ridge National Laboratory in 1980, conclude there was a single, major disaster on the basis of the ecological literature — evidence that the Los Alamos report ignores.

The Los Alamos report asserts that the incident is relevant "in the light of the current nuclear waste disposal questions facing the United States" but does not explain just how. But it implies that even when nuclear wastes were disposed of in the sloppiest manner imaginable, they caused no single catastrophe. **Deborah Shapley**

Britain in space

Up, up and away

The British government, not usually enthusiastic about spending public money, seems keen to find the cash for a major national programme in remote sensing. Members of a "task force" who have spent the past six weeks working out the details, were putting the finishing touches to their report earlier this week.

Behind the government's enthusiasm is its wish to make the best use of the European Space Agency's (ESA's) Earth Resources Satellite (ERS1) to which it has already pledged £28 million, or 13 per cent of the cost. Britain does not seem to be the only country to find an interest in remote sensing. All ESA's 13 member states are expected to meet next week's deadline for pledging their contributions to the satellite, even though it will be built under ESA's optional applications programme.

Assuming that the money is forthcoming, ERS1 will be a pre-operational ocean-monitoring satellite following along the lines of Seasat, the United States equivalent, which hinted at the usefulness of looking at the oceans in microwave before it went out of service after only three months.

Thus the three instruments that will make up the ERS1 payload — a synthetic aperture radar, a wind scatterometer and a radar altimeter — will all record in the microwave band. The satellite's low polar orbit will give it global coverage for measuring such variables as wind speed, wave height and polar ice cover. Data are expected to be of use to the scientific community, weather forecasters and shipping and offshore oil companies.

Britain's singular interest in ERS1 was stimulated last month when a team from the Science and Engineering Research Council's Rutherford Laboratory won a competition to design a fourth instrument for the payload — an infrared scanning radiometer to measure sea-surface temperature to within half a degree Kelvin.

But the instrument, costing about £1 million, must be built with British money.

If the general level of interest is anything to go by, however, the Rutherford team has little cause for anxiety. The "task force", which includes representatives from the Science and Engineering Research Council, the Natural Environment Research Council, the Meteorological Office, the Royal Aircraft Establishment at Farnborough, industry and the Department of Industry, has been drawing up plans for a far grander national remote sensing programme that could involve spending more than £4 million. The report addresses three questions: the content of a programme of fundamental research into understanding microwave data; how to get the data to the user; and how to exploit commercially remote sensing in general.

The most costly consideration could be that of disseminating data from ERS1. ESA plans to set up a central ground receiving station at Kiruna in Sweden from which data can be relayed to agencies in individual nations. But the British study describes several options for a national ground station that could receive data directly from the satellite. These range from a £4 million station capable of receiving 120 megabits per second from the synthetic aperture radar to a much cheaper station capable of handling data at no more than 1 megabit per second. The cost would almost certainly have to be met by the industry department, which may try to persuade industry itself to chip in.

The data-handling part of the national programme could ride on the tide of the British government's general enthusiasm for information technology, so that it is no surprise that the report will be placed before Kenneth Baker, minister for information technology. The other important aspect for the government, however, is how to make money out of what is widely heralded as a growing international market. One option that the industry department is considering is to persuade private industry to set up a company to sell remote sensing data, as well as hardware and software for ground stations and satellite payloads.

The final details on the form of the national programme will probably take some months to work out. In the meantime, the plans seem to beg two questions. Are remote sensing data sufficiently attractive for potential users to pay for them? And is Britain, which has no national space agency, sufficiently well organized. On the first point, the industry department seems aware that it will be treading in the dark. But an earlier survey of more than 100 potential users has persuaded it that the risk could be worth taking, especially for ocean monitoring. And the civil servants who have been debating the merits of a national space agency for some time, could be encouraged by a decision to embark on a remote sensing programme. **Judy Redfearn**

Polish universities

Warsaw's rector

Poland's Council of Ministers has resumed consideration of two bills aimed at restructuring Polish higher education, apparently undismayed that the replacement of Dr Henry Samsonowicz as rector of Warsaw University is regarded in academic circles as a confirmation of growing fears for the future of scholarly life.

Dr Samsonowicz became rector in September 1980, under a compromise arrangement between the old system of government appointment of rectors and the promised reform by which the rector would be elected by the academic community. Subsequently, he was confirmed in office by a free election in which academics, students and university auxiliary staff were represented.

Since the imposition of martial law, Samsonowicz has been under increasing pressure and, according to some sources, was temporarily detained during the round-up of intellectuals in the first few days after the military take-over. A few weeks ago, he was expelled from the Polish United Workers' Party, since when his dismissal has been thought to be only a matter of time.

Nevertheless, the announcement that Dr Samsonowicz's "resignation had been accepted" caused a major wave of protests by Warsaw academics and students. At one point, all lectures were stopped in Warsaw University for 15 minutes to allow students and staff to sign protest letters.

The new rector, selected by the Minister of Science, Higher Education and Technology, is Professor Lazimierz Albin Dobrowolski, head of the department of zoology and ecology. In spite of the unfortunate circumstances of his appointment, Professor Dobrowolski does not seem as severe a hard-liner as some pessimists have assumed.

In a discussion on the proposed reforms of academic life in November 1980, he accepted as necessary the demands for reforms at all levels of education. He even implied that, theoretically, the universities already possess the right to make changes in their syllabuses which, if true, would make many of the reforms promised by the Lodz accords of February 1981 merely a re-statement of powers which the universities already possessed but had not used.

With martial law, the reforms were cancelled and university life reverted to the statutes of 1958, which place full administrative power in the hands of the rector and reduce university senates and faculty councils to a merely consultative role. Many features of the pre-Lodz syllabuses, including compulsory Russian, have been restored.

How far the proposed new law will embody these "temporary" measures is not clear. As early as last summer, the then minister, Dr Jerzy Nawrocki, tried to

emasculate the promised bill by last minute amendments which he introduced unilaterally. At that time, the minister's action was thwarted by the threat of a nationwide university strike. Strikes, of course, are now illegal under martial law.

Vera Rich

Planetary science

Going downhill

Washington

Ever since Mr George A. Keyworth II, the President's science adviser, made some remarks to the effect that it had already had its day, planetary science has been struggling to find favour in Washington. Its trials are continuing, for its traditional patron, the National Aeronautics and Space Administration (NASA), is cutting back support, and it is not clear that the National Science Foundation (NSF), the logical alternative patron, will take up the slack. Although much more than hardware is involved, the controversy centres at the moment on the imminent closure of the relatively new Infrared Telescope Facility (IRTF) atop an extinct volcano on Mauna Kea, Hawaii.

Clearly, the entire field of astronomy is not going to get everything it wants from the government. For instance, a major study for the National Academy of Sciences (NAS) of future astronomical needs, chaired by George Field, director of the Harvard Smithsonian Center for Astrophysics (see *Nature* 8 April, p.482), has produced a list of desirable future facilities requiring \$1,700 million in expenditure over the next decade. The problem, however, is how to set priorities — a task scientists are reluctant to take on, with the result that it falls to the budget men in Washington.

Earlier this year, the Office of Management and Budget (OMB) told NASA that it could have only two of its three major efforts in astronomy. NASA opted to keep the space telescope, scheduled for launch in 1985, and the science programme for the space shuttle. This means that NASA's third major effort, planetary science, must suffer: the Earth and Planetary Sciences Division of NASA will be cut by half in 1983 compared with 1981, if the proposed \$61 million cut is not restored when Congress considers the NASA budget later this year.

The cuts affect funds that support young graduate students entering planetary science and astronomy, and small grants to astronomers around the United States, away from major facilities. Among the programmes axed was the Venus Orbiting Imaging Radar (VOIR) and funds were not restored for a US Halley's comet probe to complement the Giotto project. Moreover, although support for the solar-polar mission has risen to \$21 million, the increase supports only American work aboard the European spacecraft.

But the most conspicuous cut in the fiscal 1983 budget was the IRTF, which sits 13,000 feet above sea level on Mauna Kea. The IRTF was a pioneering facility built by NASA to observe the planets as support for Voyager and other missions. Its tasks include monitoring volcanism on Io and plotting the infrared signatures of the outer planets and their possible moons.

Axing the IRTF was not an arbitrary decision. All NASA facilities are supposed to support actual missions, and no US spacecraft will encounter a planet until Voyager reaches Uranus on 24 January 1986. Moreover, according to a policy statement made some years ago, all ground-based telescopes funded by the government should be supported by NSF, the government's basic research agency.

On 1 January 1983, the IRTF will be homeless, patronless and mothballed unless Congress restores its \$1.7 million annual operating cost in the NASA budget. NASA has recommended that NSF picks up this tab, but the NSF budget for 1983 is more or less set. Mauna Kea's administrators will have to file a proposal to NSF to get the money — but even then may well not succeed.

John T. Jeffries, director of the Institute for Astronomy at Mauna Kea (which includes other facilities besides the IRTF), is urging Congress to restore the funds to NASA. For the moment, Jeffries says he is reluctant to ask NSF for the money, which at this late stage could come only from other astronomers.

If the telescope goes, United States astronomers would have no national centre for infrared work — either within the Solar System or outside it. The infrared facilities at Wyoming and Arizona serve mostly local astronomers, while a comparable facility on Mauna Kea, the United Kingdom Infrared Telescope, is available mainly to British astronomers.

The Mauna Kea IRTF's plight is typical of the dilemma facing much of the basic science carried out by mission agencies in the United States. The facility is built and operated by NASA, so dedicates half its work in support of NASA's planetary exploration programme. But applications to work on objects outside the Solar System, such as the gas clouds of Orion, are more than double IRTF's applications for planetary work.

But whatever the merits of Mauna Kea's IRTF or the fate of Jeffries' attempts to save it, the issue not being discussed is the one Keyworth raised originally, of whether planetary science has had its day and deserves lower priority than other branches of astronomy. According to Field, chairman of the NAS study, "traditionally, NASA has tried to encourage cooperation rather than conflict between planetary research and galactic and extragalactic astronomy. It has not allowed them to come into conflict but has maintained constituencies for both."

Deborah Shapley

Biotechnology

Dutch go-ahead

Waalre, The Netherlands

The Dutch have at long last got their "innovation programme" for biotechnology, eagerly awaited since the beginning of the year (see *Nature* 14 January, p.91). A government-sponsored committee under the chairmanship of Professor R.A. Schilperoot, who works on recombinant DNA research at the University of Leiden, did not, as might have been expected, recommend setting up new centres for biotechnology along the lines of those planned in microelectronics. Instead his committee has opted for the rapid strengthening of cooperation between government institutes, universities and industry.

The committee urges that the government should provide extra support to stimulate innovation in biotechnology — at least an extra 75 million guilders between now and 1988. This money would come from the government's fund for industrial innovation and be in addition to the planned biotechnology budget, which is much larger.

Efforts should be concentrated on applied research, the committee concludes, especially in sectors where Dutch companies are traditionally strong — such as agriculture, the dairy industry, fermentation and antibiotics production. Areas highlighted for future research include the development of host-vector systems for use in applied research, somatic cell hybridization, second-generation biotechnological reactors for enzyme production and the isolation of useful products from process liquids.

Professor Schilperoot is confident that the existing large and medium-sized Dutch companies (such as Shell and Unilever) will be better able to tackle the challenge than would newly-formed small companies. Discussions on the regulation of recombinant DNA research went on in The Netherlands for far too long. Professor Schilperoot believes, badly holding up research; "The discussion should stop now. There certainly must be rules, but not stricter than necessary or stricter than in other countries. Permits to start work should be issued quickly, and the rules should be established nationally. Rules made by regional safety authorities have sometimes been too strict."

Schilperoot, who himself sometimes acts as an adviser to a foreign company, says that Dutch industry and universities often have extensive contacts abroad at the expense of national cooperation. "Better structures here would create a favourable climate that would encourage more confidence in this field and put The Netherlands in a strong position in 5 years' time and have a great impact on society and industrial activity in the next 10 to 20 years."

Casper Schuurin

Shake-up plans for National Academy of Sciences

Washington

This week, the US National Academy of Sciences (NAS) holds its annual meeting in Washington. Probably more than 300 distinguished scientists will gather. They will not transact much business of substance, but there will be much chat about the Academy's reorganization and the new vision of things of its President, Frank Press, a geophysicist who was President Jimmy Carter's Science Adviser.

The NAS is old by American standards (it was founded by President Abraham Lincoln in 1863) and has been accused of being old-fashioned, especially when it tries to give the government the definitive word on such controversial topics as marijuana use or ozone or acid rain. Clearly the academy's science panels have had difficulty giving advice when the assigned topics have been of emotional national concern. "Most of what the academy has done in the past has been extremely good and very important," says Press. "But if we do fifty reports well and one report badly, often the one report gets all the attention."

Press wants to retain NAS's reputation for fairness, impartiality and as the country's most authoritative source of science advice. He wants to continue the special relationship the NAS has long held as the government's science advisory group. But he wants reports to come out faster, review procedures to be simpler, and the NAS to exercise leadership in what topics are studied by having independent sources of income to fund self-initiated studies.

These views have already had their effect. Several key studies were initiated, either with Press's advice before or after his taking office, which are half or less fully supported by the government. Among them are:

- University research and national security headed by Dale R. Corsen, former president of Cornell University. It is due to report in a year;

- Government relations with research universities, led by Burke Marshall, professor of law at Yale University. It will study indirect costs, Circular A-21, federal regulation of the laboratory, and will report this summer;
- Relations between the industrial democracies in an era of high technology competition, led by Howard Johnson, Chairman of the corporation, MIT. It is due to report by the end of 1982;
- Arms control and international security, led by Marvin Goldberger, President of California Institute of Technology;
- Ageing. This study is just getting under way. "The government wasn't interested in supporting this" says Press.

Press has started to raise money for NASA's endowment, most of which was raised by former president Frederick Seitz and which now stands at about \$30 million. Three foundations have given \$3 million so far and a fourth, the Alfred P Sloan Foundation, has agreed instead to help fund the government-research university-study. Meanwhile, NAS's sister organization, the National Academy of Engineering and the Institute of Medicine (IOM), are starting fund-raising drives of their own.

Press is also planning to seek funds from industry through a corporate affiliate programme. "People will say how can you separate donations from control of the work" Press said, anticipating criticism of this move. "We plan to have companies put their membership fees in a blind pool, which we could then draw on without corporate advice to fund our work." In return, member corporations would be invited to special symposia and get all academy reports.

Press, conscious of criticism that academy reports come out slowly and that review procedures are cumbersome, intends to have a special management procedure for the 20 or 30 reports that are nationally significant, thus ensuring that they are "timely" and "dependable".

The studies could be revised faster — "a review should not take more than 2 or 3 months" says Press. Usually, the board overseeing a study does one review of it; then it goes through an separate review procedure run by the National Research Council, NAS's research arm. Press is experimenting with having both reviews go on at once or having the board waive its review if it has a say in the review run by NRC. Another way to speed up reports is not to insist on consensus, but to let dissenters express their views separately. "I've encouraged that", Press says.

To "simplify" NAS procedures, Press is abolishing all the assemblies set up by Handler ten years ago, and two of the future commissions. An *ad hoc* panel under James Ebert, president of the Carnegie Institution of Washington, looked

at the academy structure and recommended changes, most of which have been implemented by NRC in the past month or so (see *Nature* 1 April p.385).

Under the "streamlined" NAS organisation the old established offices such as the Transportation Research Board and the Office of International Affairs will become their own bosses again, their parent commissions abolished or merged with other things. Agriculture is given new priority, and will probably become its own research board. More attention will be paid to primary and secondary level technical education by the Commission on Behavioural and Social Sciences and Education. What will happen to the



Frank Press - businesslike approach

Institute of Medicine and the Assembly of Life Sciences is uncertain, although there will probably be a new life sciences commission, excluding agriculture. Oceans activities will be consolidated from four boards into two.

Press can implement these changes without a vote of the membership and indeed has already done so. One reason he is seeking non-government money, he says, is to stabilize the staff situation; he plans to hold more regional meetings of members. While most members have accepted the changes, and probably welcome the additional outside funds, Press's businesslike approach to the academy has aroused some murmurs — especially from people in commissions being abolished.

Deborah Shapley

"DON'T TALK TO ME ABOUT AGEING!"



Correction

Professor E.A. Barnard, not Professor B. Hartley, is head of the biochemistry department at Imperial College, London (*Nature*, 1 April page 382). *Nature* apologises to both professors.

CORRESPONDENCE

University role

SIR — Your analysis in *Science in France* (*Nature* 25 March, pp.285–304) prompts me to respond, especially on the subject of the role of the universities in the research effort.

First, your articles clearly transmit the French media's impression that the best research in the country is done outside the universities in the *grand organismes* research centres. Nothing could be further from the truth: many of the most prestigious groups of CNRS and INSERM are directed by university professors (such as P. Chambon, J. Dausset, J.F. Bach, M. Seligmann . . .), and they are also closely involved in the training of future investigators.

Many CNRS and INSERM laboratories are actually located in university campuses, and research teams usually include both university staff and CNRS and INSERM appointees. In addition, many of the CNRS, CEA and INSERM laboratories located outside of the campuses are staffed in part by university faculty. Thus at the Institut Pasteur or at Institut de Recherche en Biologie Moléculaire from the CNRS, a number of teams are directed by university professors, actively involved in the training of the students of the various investigators.

A second point arising from your articles is the future role of the universities in research. The faculties of sciences have been heavily involved in both research and training in the past and will continue to be so. This does not go without a rigorous selection of potential candidates, a fact which remains generally unnoticed by the media: our own "UER de Biochimie" (Unit for Education and Research in Biochemistry) has about 700 undergraduate students: only 70 are admitted to the graduate courses in microbiology, virology, immunology, molecular biology and nutrition. The first three of these courses are organized by the University of Paris's staff on the premises of the Institut Pasteur.

Incidentally, most of the graduates will end up in CNRS and INSERM research, not only because no new openings are available at the universities but also because these students are the best trained in the field — a fact which we would like our new ministers to be aware of.

A. DONNY STROSBURG

CNRS Institut de Recherche en
Biologie Moléculaire,
Université Paris VII, Paris, France

Mosquito war games

SIR — In his report (*Nature* 11 March, p. 104) about allegations of American biological warfare research in India and Pakistan, Dr K.S. Jayaraman is altogether too modest in omitting mention of his own primary role in the press campaign on this topic in India in 1974–75. That campaign was long ago shown to have been based on incorrect information and erroneous reasoning (*Nature* 252, 342; 256, 355; 258, 102; 273, 96). However, as a result of the atmosphere of suspicion created by that campaign all sorts of difficulties are still put in the way of international collaborative work in India on topics of vital importance to the Indian people, such as nutrition and malaria.

Jayaraman correctly states that yellow fever

does not occur in India and he suggests that the only reason to study the "yellow fever mosquito", *Aedes aegypti*, would be in support of biological warfare. However, as he was told in 1974, *Ae. aegypti* is also the vector of dengue and dengue haemorrhagic fever in India and elsewhere. He was always inclined to belittle the importance of these diseases but certainly *Literaturnaya Gazeta* does not do so, and one of its more bizarre claims was that mosquitoes were being reared in Pakistan and transported to the Caribbean area for use in biological warfare against Cuba, using dengue virus.

It has been claimed that work was planned or carried out at the Pakistan Medical Research Centre (PMRC) on genetic manipulation of *Ae. aegypti* and on Japanese encephalitis virus. However, Professor R. Baker, ex-director of PMRC, the Annual Reports of PMRC and our own research visits there make it quite clear to us that this was not so. A small amount of work was done on West Nile virus which is indigenous to Pakistan. This work was to study the feasibility of producing a mosquito strain not susceptible to transmission of this virus. Such work may lead to vector control by the introduction of non-susceptibility genes into wild mosquito populations. The main work of PMRC was on methodology for *Anopheles* and *Culex* control by release of semi-sterile males and several elegant genetic sex separation systems were developed to ensure that only males were released. Male mosquitoes do not bite and therefore could not possibly be the agents of biological warfare.

PMRC has produced a remarkable amount of excellent published work. It would be a shame if the history of the WHO/ICMR unit in Delhi should be repeated and the work of PMRC were hampered or terminated as a result of ill-informed and prejudiced journalism.

G. DAVIDSON, C.F. CURTIS
G.B. WHITE & P. RAWLINGS

London School of Hygiene
and Tropical Medicine,
London WC1
and Tsetse Research Laboratory
University of Bristol, UK

Propaganda wars

SIR — Yet again a leading article your journal (1 April, p.380) disparages the work of Scientists Against Nuclear Arms, the Medical Campaign Against Nuclear Weapons and other groups of professional people dedicated to reducing the present dangerous accumulation of nuclear weapons. The charges brought against these groups seem to be the following: They can only partially address the problem and offer no practical advice; they claim a spurious unity among the profession concerned; they are prone to oversimplifying the issues; and finally, their influence is somehow "irrational". These charges are no more than risks (some more real than others) to be avoided. They are not sufficient reason to condemn the existence of these groups. Perhaps a paragraph's explanation may allay fears that the emergence of these groups is unhelpful or even dangerous.

Education is obviously a place where the informed should act — not individually but in

groups which can bring the information to the public more effectively. This is made even more necessary by the fact that the government's own literature on nuclear questions seems bent on misinformation and deception. As scientists, we have an especial responsibility to inform the public not only because of our knowledge of relevant scientific matters but also because of our understanding of the armaments industry and the complex ways in which technology, warfare and politics interact. Thus, aware of the dangers inherent in the present situation, we need effective representation of our views. This is provided by the professional groups which are now being formed. Moreover, it is essential to the disarmament movement that it is seen to be as widely based as possible, including professional people as well as parents, Christians, politicians and so on. Diversity here is strength, not weakness.

There are many of us, well informed on the present situation, who are convinced that we shall die within the next few years in a nuclear conflict unless the course of current events can be turned. There are a few signs of hope however, one of which is the emergence, rapid growth and energetic activity of the professional groups that you misguidedly choose to malign. I ask that your journal supports the work of these groups so that mankind can again be confident of its future.

D.E. RYDEHEARD
A. MYCROFT

Edinburgh University,
Edinburgh, UK

SIR — Your editorial of 1 April ("Professional propaganda", p.380) concludes that groups such as Scientists Against Nuclear Arms, Teachers for Peace and the Medical Campaign Against Nuclear Weapons are deceitful and intrinsically undesirable because they foster the false impression that their professions are united on the issue of nuclear weapons, and because they do not address the relevant issues. This is a contentious inference based on dubious assertions, and it is unusual for *Nature* to adopt a deliberately provocative and weakly argued position so firmly. This political use of your unique position in scientific journalism seems to me a more egregious abuse of professional influence than the overtly political activities of the scientists and other groups that you have attacked in this way.

JULIAN PETO

The Radcliffe Infirmary,
Oxford, UK

On the flat

SIR — M. Hammerton (*Nature* 18 March, p.192) asks why you continued so long to publish letters on creationism, arguing that you would not give such space to flat-Earthers, for instance. But on the same page is a long letter from Professor J. W. Jeffery, citing, as the core of his argument, an unpublished study of his own and a 1979 work of sci-fi and crystal-gazing in energy technology as if it were a work of economic fact.

Surely it is time we heard from the flat-Earthers.

DAVID FISHLOCK

Financial Times,
London EC4, UK.

NEWS AND VIEWS

Lead isotopes and the Bronze age metal trade

from Keith Branigan

FOR the past twenty years archaeological scientists have spent much time and effort trying to trace the metals used by early man back to the ore sources from which they were originally recovered. Attention has been concentrated on the possibilities of identifying characteristic patterning of trace elements in the metal. Research published within the last twelve months by Noël and Zofia Gale of Oxford University, however, has shown that successful provenancing of metals is now possible — by using lead isotope analysis. Preliminary results have already produced some major surprises about the supply of metals to the civilisations of the Aegean Bronze Age (~2000–1200 BC).

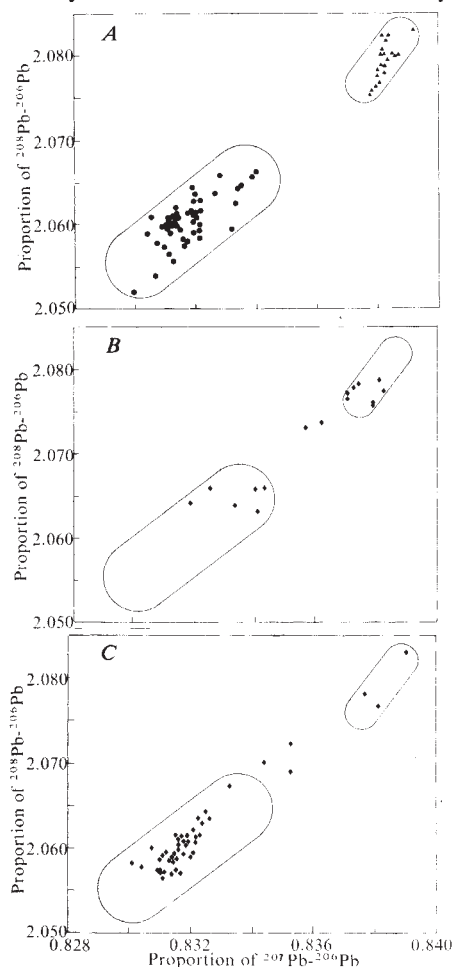
Despite the production of many thousands of trace element analyses, both archaeologists and archaeometallurgists have remained sceptical of the successful application of this method to the provenancing of ancient metals. The chemical composition of ores seems to vary within the same ore body, and early smelting techniques resulted in poorly controlled changes in concentrations of trace elements in the metal they produced. An alternative method has therefore been sought by Noël and Zofia Gale, working from the Department of Geology and Mineralogy at Oxford. Their studies have concentrated on lead isotopes which have two important characteristics. First, the lead isotope composition of different lead ores differs according to their geological age and the amounts of uranium and thorium initially present. Thus a lead ore body can be 'finger-printed' by isotope analysis. The second significant characteristic of lead isotopes is that their composition pattern is not altered by the metallurgical processes. This means that not only ores, but also objects made of lead, silver, or other metals containing lead can also be 'finger-printed', and related directly to an ore source. Five years of research enabled the Gales to publish the 'finger-prints' of various Aegean lead-silver sources in June 1981 (*Scient. Am.* **244**, no. 6, 176), and to announce the isotope analyses of about 90 Bronze Age

artefacts of lead and silver from the same region. The analyses of ores revealed that lead-silver ores from the islands of Siphnos, Antiparos and Syros, and from the site of Laurion in Attica could all be clearly distinguished by their varying isotope composition patterns (Fig. A). The analyses of artefacts immediately

produced a major surprise — the importance of the Laurion lead-silver source in early Aegean metallurgy. Its significance to classical Athens is well known and documented, but no one had believed that its metals had been extensively used 1,000 years earlier. In fact, the Gales' analyses showed that even in the early Bronze Age — 2,000 years before classical Athens — the Laurion mines supplied silver and lead to the islands of the Cyclades, the rest of Cycladic silver and lead coming mainly from Siphnos (Fig. B). In the second millennium BC, the Siphnos source was scarcely used and lead and silver throughout the Cyclades and Crete, as well as on the Greek mainland, came mainly from Laurion (Fig. C). But Laurion silver and lead travelled even further afield, for among the artefacts analysed by the Gales were six from Egypt, all of which proved to be made of the distinctive Laurion lead and silver. The earliest of these items dates to the 10th Dynasty, about 2100 BC, and supports other archaeological evidence which points to an opening-up of Aegean trade with the east Mediterranean in the centuries either side of 2000 BC.

At the 5th Creteological Congress in September 1981, at Ayios Nikolaos, Crete, the Gales presented two further papers (not yet published). Zofia Gale described new isotope analyses for lead-silver ores from Crete (characterized in each case by a very high $^{207}\text{Pb}/^{206}\text{Pb}$ ratio) and also some 40 new analyses of Minoan lead-silver artefacts. They showed that whilst most of the early Bronze Age items in Crete were made of Siphnian lead/silver, a few items of Laurion metal were already reaching the island. Again, this supports the evidence of imported mainland pottery at Knossos around 2500 BC, and of Minoan-style (if not Minoan-exported) trinkets at Ayios Kosmas in Attica at the same period. One major controversy remained unsolved by the analyses — the origin of three silver daggers found at Koumasa in Crete, and variously claimed as Italian, Cycladic and Minoan in manufacture. Zofia Gale's analysis of one of these daggers showed it was made of neither Siphnian nor Laurion metal; its origins cannot yet be determined.

Noël Gale's paper to the Congress and a newly published article in *Science* (**216**, 11; 1982) are potentially the most significant



A, isotope analyses of lead ores from Siphnos (triangles) and Laurion (circles) cluster in two quite distinct groupings (simplified from Gale & Gale 1981). B, lead isotope analyses of 16 lead artefacts from the Cyclades, ~2500 BC, show 8 were made of Siphnian lead, 6 of lead from Laurion and 2 cannot yet be provenanced. C, analyses of 46 lead/silver objects from Crete and Thera (Cyclades) show that by ~1500 BC, almost all objects were made of lead from Laurion on the Greek mainland (field at bottom left).

Keith Branigan is in the Department of Prehistory and Archaeology, University of Sheffield, Sheffield S10 2TN.

for archaeology world wide, however, for they present the first analyses of copper and bronze artefacts by the lead isotope method. Copper and bronze were used for far more artefacts in antiquity than was lead or silver, and in much of prehistoric Europe, for example, lead and silver were scarcely known. Gale's demonstration that the amounts of lead found as an accidental inclusion in coppers and bronzes is sufficient for isotope analysis opens up wide possibilities of identifying prehistoric metal trades with confidence for the first time. Copper sources used by the Minoan bronzesmiths have been the subject of much debate, and Italy, Crete and Cyprus have all been proposed, with Cyprus the undoubted favourite in the late Bronze Age. Gale has analysed copper ores from Crete, Cyprus and Laurion and established their isotope composition patterns. The first batch of 22 copper/bronze artefacts analysed, all from Crete, surprisingly suggested that there were perhaps four

different ore bodies producing the metal from which these objects were made, although the location of only one of them can yet be fixed — and that is Laurion! Half of the objects analysed come from Laurion, which has never previously been thought of as a potential source of copper for the Aegean civilisations. None of the artefacts was definitely of Cypriot origin, and this result, preliminary as it may be, must raise afresh the old controversy over the famous ox-hide ingots of copper. Found in Cyprus, Crete, Turkey, central Greece, Sicily and Sardinia, and shown in Egyptian tomb paintings, these heavy slabs of copper clearly represented a major Bronze Age trade in metal. But did they go from the west Mediterranean to the east, or from the east to the west, or from the centre to both west and east? Most favoured a Cypriot origin, but Gale's analyses carry at least one clear message — we must think again, and await more analyses, before we can finally solve the problem. □

However, Emlen³ goes on to show that there are often conflicts of interest in these breeding groups and he defines the precise circumstances when these will occur. Sometimes the young will increase their chances of survival by staying on the parental territory but the parents will suffer because of increased competition for resources. In this case the young could decrease the costs of their presence by helping as 'payment' for permission to stay. Breeder-helper conflict will be intensified when r between the helper and the nestlings is low (for example, because one or both of the breeders have been replaced). When there are few outside vacancies helpers may be expected to challenge breeders for breeding rights in the group. There is good evidence for this kind of conflict in pied kingfishers, *Ceryle rudis*, where some individuals offer to feed the breeders' offspring with fish but are often chased away. Here the helpers (who are unrelated to the breeders) may be offering food as a payment for access to the breeding female as a potential mate⁷.

In theory, conflict can also occur in the opposite direction. Emlen suggests that this will be especially prevalent in fluctuating environments where, when conditions for breeding are good, the helpers do better by leaving home. From their parents' point of view, however, they only rear grandchildren ($r=0.25$) if they leave, whereas if they stay they help the parents to raise more children ($r=0.5$). Forceable retention of helpers may be impossible but the breeders could, in principle, manipulate circumstances in their favour by disrupting the breeding efforts of their dispersed young so that they came back home to help instead. Trail, Strahl and Brown⁸ report destruction of nests and eggs and the murder of nestlings in groups of Mexican Jays, *Aphelocoma ultramarina*. In these examples, the murderers were unrelated to their victims but the effect is similar to that envisaged by Emlen. Helpers in the Jay groups may feed the young in several nests at once and so a possible benefit of the infanticide is that it increases the number of helpers available for the killer's own nesting attempt.

Emlen provides an elegant explanation for why there is both conflict and cooperation in breeding groups and the behavioural evidence is in qualitative agreement with the theory. To test the model quantitatively, however, we need to be able to calculate the precise reproductive payoffs to breeder and helper if the helper stayed or if it left. All these calculations to date have relied on measurements of natural variation in the population. For example, if groups with helpers have greater success than groups without, the difference is regarded as the helpers' contribution. In a similar vein, the reproductive success of a novice breeder is regarded as a measure of what a helper would achieve if it left home and bred. The problem of using observations on natural

Cooperation and conflict in breeding groups

from N. B. Davies

IN over one hundred and fifty species of birds, some individuals forgo breeding and, instead, assist others to feed and care for offspring. Such 'helping at the nest' is as an example of 'altruism' because the helpers increase the reproductive success of the breeders at a cost to themselves. Field studies of marked individuals have shown that, in many cases, the helpers are previous offspring of the breeders and stay at home to help their parents raise more broods. Because the young in the nest are close relatives (often full sibs) a helper does increase its gene contribution to future generations through helping. But why doesn't it adopt the alternative option of leaving home and breeding?

Charnov¹ has pointed out that most birds have biparental care and two individuals, a disperser plus its mate, could normally rear more young than one alone. By leaving home and pairing up a bird could raise more offspring (coefficient of relatedness, $r=0.5$) than it could by staying alone at home and helping to raise extra full sibs (where r is again 0.5). If it stays at home a helper thus forfeits half of its potential brood — the offspring its mate would have reared. Are there, then, ecological conditions which might force young birds to stay at home? Two recent papers by Emlen^{2,3} examine the ecological correlates of helping in birds and provide a model for the evolution of social behaviour in breeding groups.

Helping occurs in a wide variety of environmental conditions. Brown⁴ was the first to point out that individuals are more likely to remain on their parental territory

in very stable habitats where populations have expanded to the carrying capacity of the environment so that very few breeding vacancies exist. Good evidence comes from studies of the acorn woodpecker, *Melanerpes formicivorus*. In coastal California the habitat is extremely saturated and during a three year study⁵ no breeding vacancies arose. On average there were five adults per territory because many young had stayed at home to help their parents. By contrast, in the more seasonal Huachuca mountains of south-eastern Arizona there were many vacant territories and most young dispersed and bred⁶.

Helping at the nest also occurs at the opposite ecological extreme, in very harsh and fluctuating environments, such as the arid areas of Australia and Africa, where rainfall, and hence food supply, is very variable and unpredictable. Emlen² suggests that during harsh periods the costs of leaving to rear young are prohibitive for novice breeders and so many stay at home until conditions improve. Measurements show that young white-fronted bee-eaters, *Merops bullockoides*, stay at home and help their parents when the insect food supply is low, but in times of plenty they leave and breed themselves.

At first sight it would seem that helping at the nest should be a truly cooperative affair; the breeders gain through help with their brood while the helpers increase their inclusive fitness through raising siblings.

N. B. Davies is a Demonstrator in Zoology at the University of Cambridge and Fellow of Pembroke College.

variation to quantify the parameters in the model is one of confounding variables. Helpers which leave home and breed may be better-quality individuals than those that stay at home. Pairs with helpers may be better-quality breeders than those without; indeed, if the helpers are previous offspring, then their presence may be a consequence of the breeders' high reproductive success in the past rather than a cause of present success.

Experimental manipulations are needed to quantify the helper's contribution and the payoffs it would achieve by leaving home. Field workers are understandably reluctant to carry out large-scale manipulations of their study populations because this interferes with the long-term collection of data on natural success. Brown and his co-workers⁹ have now reported the first removal experiment of helpers in the grey-crowned babbler, *Pomatostomus temporalis*. Groups with helpers removed did worse than control groups, showing conclusively that the helpers themselves increase reproductive success.

Further experiments of this kind are needed, particularly as different explanations are likely to emerge in different situations. An intriguing case is that of the dunnoek, *Prunella modularis*, where pairs plus a helper have lower reproductive success than pairs without¹⁰. Before we jump to the conclusion that the 'helpers' must in fact be hindering the pair, a removal experiment is needed. An alternative possibility is that helpers only manage to force themselves on to territories with poor-quality breeders; the breeders may have done even worse without the helper!

Finally, it is interesting to note that cooperative breeding often occurs in groups of unrelated individuals. In the Mexican Jay¹¹, unrelated immigrants help feed other pairs' nestlings and defend the territory in the same way as close relatives. With all the recent emphasis on the role of kin selection in the evolution of social behaviour we may have underestimated the importance of mutualism and reciprocity (see ref. 12). In this issue of *Nature* (p.740), Packer and Pusey show that there is just as much cooperation in unrelated coalitions of male lions as in related groups. Although a male would increase its inclusive fitness even more by cooperating with a brother, it still pays to cooperate with non-relatives simply because large coalitions father more surviving young per male than small coalitions. □

Will the real primary magma please stand up?

from J.M. Rhodes

If basaltic magma, produced by partial fusion of the Earth's mantle, were to rise through the mantle and overlying crust, ultimately to be erupted from some volcano with no change in composition, it would be avidly sought and studied by petrologists and geochemists intent on deciphering the composition and mineralogy of the mantle. Basaltic magmas such as these, which have made their way to the Earth's surface without change in composition, are usually referred to as 'primary magmas' since they are the parental compositions from which other, more evolved magmas are produced by a variety of processes as the magma cools, crystallizes and ultimately solidifies. They are regarded as a sort of Rosetta Stone to the Earth's interior, and, indeed, to the interior of any planet, providing suitable samples are available.

However, there is little agreement on whether a primary basalt has ever been sampled, and their nature and composition remain subjects of spirited debate even after almost twenty years. With the advent of the plate-tectonics hypothesis, basalts that have been erupted along the vast 65,000 km interconnecting, spreading mid-ocean ridge systems have figured prominently in this debate, and discussion has mirrored the arguments concerning primary magmas in general.

On the one hand, typical mid-ocean ridge basalts are considered the best candidates for primary magmas because of their vast volumetric abundance, which exceeds the products of more conspicuous forms of volcanism by an order of magnitude, coupled with an apparent world-wide compositional uniformity¹⁻³. Alternatively, and currently favoured, is the argument that mid-ocean ridge basalts, and indeed almost all basalts, are far removed from primary compositions, being instead the end products of an extensive and complex differentiation process from a primary picritic magma; a process that may include crystallization, contamination and mixing with other magmas, as the magma makes its way to the surface through the fractures, conduits and chambers of the volcanic plumbing system⁴⁻⁶. Somewhere between these two extremes is the view that only the most primitive basalts sampled from the ocean floor are primary⁷⁻¹⁰.

All these interpretations have their difficulties. Most mid-ocean ridge basalts show both mineralogical and chemical evidence for crystallization at shallow levels within the crust, an observation clearly at odds with a primary status. Objections to the primary nature of 'primitive' basalts come from melting

experiments, which show that their compositions are at variance with the picritic melt compositions predicted to form at depths of over 30 km in the mantle. The absence of picrites amongst samples recovered from the ocean floor is a major obstacle to their acceptance as primary magmas, and has led to a number of cunning models which attempt to explain why these picritic magmas cannot rise through either the crust or a pre-existing magma chamber¹¹⁻¹³.

These difficulties, and those of identifying primary magmas and understanding the processes that relate them to the commonly erupted mid-ocean ridge basalts, are addressed in a recent comprehensive paper by Bryan and co-workers¹⁴ dealing with the compositional variations of basalts erupted along the Mid-Atlantic Ridge adjacent to the Kane Fracture Zone. In contrast to most papers on the subject, Bryan *et al.* argue for a spectrum of primary magma compositions, rather than a single magma type from which all other ocean-floor basalts are derived. They identify at least two distinct compositions that are parental to most other basalts in the Kane Fracture Zone region through complex interaction of crystallization and magma-mixing processes^{10,11,15}. One of these parental compositions is not unlike the proposed 'primitive' basalt primary magma. Although not abundant, it can be recognized as an end-member component in a number of mixed magma compositions. The other, more widely distributed parental magma does not have a 'primitive' composition, and is, in fact, close to estimates of average mid-ocean ridge basalt compositions¹⁶.

Of these two parental magmas Bryan *et al.* believe that the 'normal' rather than the 'primitive' variety is the more appropriate candidate for a primary magma, thus contrasting with the recent preference for elusive, if not illusory, primary compositions, and returning to the idea that primary magmas should be abundant and widely distributed, and parental to the less abundant, evolved basalts associated with them¹⁷. They point out that although this suggestion may appear at odds with chemical criteria commonly used to identify primary magmas, it is not contradicted by evidence from melting experiments^{6,18,19}, providing the magmas formed at shallow levels in the mantle. The 'primitive' parental basalts, on the other

1. Charnov *Anim. Behav.* **29**, 631 (1981).
2. Emilen *Am. Nat.* **119**, 29 (1982).
3. Emilen *Am. Nat.* **119**, 40 (1982).
4. Brown *Am. Zool.* **14**, 63 (1974).
5. MacRoberts & MacRoberts *Ornithol. Monogr.* **21**, (1976).
6. Stacey & Bock *Science* **202**, 1298 (1978).
7. Reyer *Behav. Ecol. Sociobiol.* **6**, 219 (1980).
8. Trail, Stranhl & Brown *Am. Nat.* **118**, 72 (1981).
9. Brown *et al. Science* **215**, 421 (1982).
10. Birkhead *Ibis* **123**, 73 (1981).
11. Brown & Brown *Science* **211**, 959 (1981).
12. May *Nature, News and Views* **292**, 291 (1981).

J.M. Rhodes is Five College Associate Professor of Geochemistry in the Department of Geology and Geography, University of Massachusetts, Amherst, Massachusetts 01003.

hand, are much less abundant, and do not correspond in composition with experimentally predicted mantle melts produced at any depth. Bryan *et al.* suggest that these 'primitive' basalts are not primary, but are the differentiates of the elusive primary picritic magmas.

Thus, they envision a wide range of possible primary magma compositions, depending on the depth of formation and segregation from the mantle. The corollary to this is that compositionally similar basalts may have been produced in quite different ways, either by direct melting of the mantle or through differentiation of a more primitive melt *en route* to the surface. This last point may explain the apparent paradox that while compositional criteria require massive amounts of crystallization to produce evolved basalts from 'primitive' or picritic magmas²⁰, phenocryst assemblages and melting experiments commonly point to a much less extensive process. However, others have explained this same paradox as a consequence of mixing 'primitive' magmas with more evolved ones in crustal magma chambers^{15,21}.

In keeping with the controversial nature of the subject, it is unlikely that the scheme presented by Bryan and co-workers for primary basaltic magmas from the Kane Fracture Zone region will win universal acclaim. If correct, it has several important implications. Perhaps the most immediately disquieting, particularly to geochemists, is the realization that the commonly used compositional criteria for identifying potential primary magmas may be of little value. More far-reaching, however, and an aspect not explored by Bryan *et al.*, is the implication that if some of the 'normal', widely distributed ocean-floor basalts are primary and not evolved, then substantial portions of the sub-

oceanic mantle must be a more fertile source of melts, and richer in iron and incompatible trace elements, than popular mantle models suggest. Whatever one's prejudices, it is clear that the concept of multiple primary magmas will receive more attention in the future. In the paper under discussion the authors argue for primary magmas that are either picritic or widely distributed normal mid-ocean basalts as a consequence of melting at various depths in the mantle. Others appeal to schemes of

incremental melting of increasingly refractory mantle in an ascending diapir to produce primary magmas with differing compositions^{8,22,23}. To resolve these increasingly complex concepts, careful and comprehensive major, trace, isotopic, petrographical and experimental melting studies of samples from well documented basaltic terrains will be needed. If these are forthcoming, perhaps the primary magma controversy will have disappeared after the next twenty years. □

Interferon redux

from Robert M. Friedman, Lois B. Epstein and Thomas C. Merigan

AN international symposium on the chemistry and biology of the interferons*, held a few weeks ago, saw rapid progress reported in four main areas of investigation: structure and function of interferons (IFNs) and their genes, animal and *in vitro* models for IFN action, IFNs in human diseases, and clinical trials with recombinant and natural IFNs.

Charles Weissman (University of Zurich) reported that 14 different genes for human IFN- α are now known, 3 pseudogenes have been sequenced, and 4 others are known that hybridize but, as yet, have not been sequenced. He has now cloned mouse IFN- α , and suggests that, whereas the genes for IFN- α and IFN- β seem to have diverged some 400 million years ago, those for human and mouse IFN- α diverged about 70 million years ago. D. Goeddel (Genentech Inc.) has cloned four genes for bovine IFN- α and he and his colleagues have also cloned IFN- γ in *Escherichia coli* and monkey cells and achieved its cloning and partial secretion in yeast. The gene has three introns and four exons and is located on human chromosome 12.

Natural IFN- γ has been purified to homogeneity (J. Vilcek, New York University School of Medicine) and is found in 20 and 25,000 molecular weight forms. ¹²⁵I-labelled IFN- γ binding studies show the receptor for IFN- γ and IFN- β to be the same, but distinct from that for IFN- α . In preliminary antiproliferative studies on HeLa cells, pure IFN- γ was no more potent than IFN- α . High-level production of cloned IFN- β (M. Innis, Cetus, Berkeley) is associated with distinct morphological changes in the producing bacteria.

Recombinant IFN- α has been crystallized by Weissman's group and by S. Pestka (Roche Institute) and his colleagues, but X-ray diffraction studies have yet to be used to elucidate the tertiary structure of the molecule. From the location of hydrophobic and hydrophilic

residues, R. Wetzel (Genentech Inc.) predicted that IFN- α is a globular protein with a hydrophobic core, more than 65 per cent of which is in α -helical formation connected with intervening β -sheets. Both he and Pestka indicated that the last 11 amino acids of the recombinant A IFN are not necessary for its antiviral effect; in fact, Pestka and his colleagues have found three natural IFN- α molecules that are only 155 amino acids long, compared with most other IFN- α molecules which have 165 or 166 residues. Pestka argued that the potency of various IFNs should be defined in terms of the number of molecules necessary for achieving a given effect now that the availability of pure recombinant IFNs makes this possible. His and others' data make clear that various recombinant IFNs have different degrees of antiviral and antiproliferative activity when compared on a per molecule basis. Despite the differences, L. Epstein (University of California, San Francisco) showed that natural IFN- α and recombinant IFN- α (A) induce the same peptides in normal fibroblasts. Recombinant, like natural, IFN- α also induces 1.5–2 times as much of these peptides in fibroblasts trisomic for chromosome 21.

Several groups have inserted the genes for human IFN- β in murine cells and J. Collins and his colleagues (Gesellschaft für Biotechnologie) have found that at least five other genes are co-regulated with the IFN- β gene, but activation of these genes is not necessary for production of IFN- β in this heterologous background.

The antiviral and cell growth inhibitory effects of IFNs can be separated in a Swiss mouse cell clone that remains sensitive to the former, but has lost sensitivity to the latter (T. Sreevalsan, Georgetown University). IFN treatment was also shown to result in a decrease in ADP

Robert M. Friedman is in the Uniformed Services University of the Health Sciences, Bethesda, Maryland 20814; Lois B. Epstein in the University of California, San Francisco, California 94143; and Thomas C. Merigan in Stanford University, Stanford, California 94305.

- Engel, A.E.J. *et al.* *Bull. Geol. Soc. Am.* **76**, 719 (1965).
- Green, T.H. *et al.* *Earth planet. Sci. Lett.* **2**, 41 (1967).
- Carmichael, A.S.E. *et al.* *Igneous Petrology* (McGraw-Hill, New York, 1974).
- O'Hara, M.J. *Nature* **220**, 683 (1968).
- Yoder, H.S. *Generation of Basaltic Magma* (National Academy of Sciences, Washington DC, 1976).
- Stolper, E. *Contr. Miner. Petrol.* **74**, 7 (1980).
- Frey, F.A. *et al.* *J. geophys. Res.* **79**, 5507 (1974).
- Langmuir, C.H. *et al.* *Earth planet. Sci. Lett.* **36**, 133 (1977).
- Bryan, W.B. *J. Petrol.* **20**, 293 (1979).
- Rhodes, J.M. *et al.* *Tectonophysics* **55**, 35 (1979).
- O'Hara, M.J. *Nature* **206**, 503 (1977).
- Huppert, H.E. & Sparks, R.S.J. *Contr. Miner. Petrol.* **75**, 279 (1980).
- Stolper, E. & Walker, D. *Contr. Miner. Petrol.* **74**, 7 (1980).
- Bryan, W.B. *et al.* *J. geophys. Res.* **86**, 815 (1981).
- Walker, D. *et al.* *Contr. Miner. Petrol.* **70**, 111 (1979).
- Melson, W.G. *et al.* in *The Geophysics of the Pacific Ocean Basin and its Margins*, 351 (American Geophysical Union, Washington DC, 1976).
- Turner, F.J. & Verhoogen, J. *Igneous and Metamorphic Petrology* (McGraw-Hill, New York, 1960).
- Presnall, D.C. *et al.* *J. Petrol.* **20**, 3 (1979).
- Fujii, T. & Kushiro, I. *Yb. Carnegie Instn Wash.* **76**, 401 (1977).
- Clague, D.A. & Bunch, T.E. *J. geophys. Res.* **81**, 4247 (1976).
- Rhodes, J.M. & Dungan, M.A. in *Deep Drilling Results in the Atlantic Ocean: Ocean Crust*, 262 (American Geophysical Union, Washington DC, 1979).
- Wood, D.A. *Geochim. cosmochim. Acta* **43**, 1031 (1979).
- Basaltic Volcanism Study Project *Basaltic Volcanism on the Terrestrial Planets*, Sec. 1.4.2, 409 (Pergamon, New York, 1981).

*The symposium on "Chemistry and Biology of Interferons: Relationship to Therapeutics" was held on 7–12 March 1982 at Squaw Valley, California. It was one of the 11th Annual UCLA Symposium on Molecular & Cellular Biology to be published by Academic Press, New York.

ribosylation and in protein methylation. There was much discussion of the role of the IFN-induced protein kinase and 2',5'-oligoadenylate (2',5'-A_n)-related systems in the actions of IFNs. The 2',5'-A_n-activated endonuclease is important in determining the antiviral action of IFNs (I. Kerr, Imperial Cancer Research Fund). The endonuclease may also be important in IFN inhibition of cell growth; endonuclease activity was absent in human melanoma cells resistant to growth regulation by IFN, but present in melanoma lines sensitive to it (A. Creasey and T. Merigan, Stanford School of Medicine). The 2',5'-A_n system may have other roles: the degree of natural killer cell activity in large granular lymphocytes was increased by transfection with 2',5'-A_n (R. Herberman, NIH). P. Torrence (NIH) explored the significance of the IFN-induced kinase in inhibiting *in vitro* viral protein synthesis. Adding the double-stranded RNA, poly(A) poly(2'-fluoro, 2'-deoxyuridylic acid), that activates the protein kinase, but not the 2',5'-synthetase, to extracts of L cells does not cause inhibition of protein synthesis; although the same inducer inhibits protein synthesis in reticulocyte lysates. Several other model systems for IFN activity were presented. In mouse cells transformed by bovine papillomavirus, in which the viral DNA remained as an episome, the viral DNA copy number per cell was decreased by IFN treatment (P. Howley, NCI, and R. Friedman). After 10 passages morphological cell revertants, entirely lacking in viral DNA, were found. Such 'cured' cells could be again transformed by the virus.

IFNs were reported to be present in a variety of human diseases, but their origin and effects remain cryptic. An unusual species of acid-labile IFN- α is present in about 50 per cent of patients with systemic lupus erythematosus (SLE) and in male homosexual patients with Kaposi's sarcoma or lymphadenopathy (O. Preble, J.A. Sonnabend (NIH), R. Friedman & J. Vilcek). One SLE patient was found to carry antibody to human IFN- α . The glomerular basement membranes from the kidney of another SLE patient with membranous glomerulonephritis developed fluorescence when sequentially exposed to rabbit anti-human IFN- α and to fluorescent anti-rabbit IgG (S. Panem, University of Chicago). The latter suggested that immune complexes containing human IFN- α may contribute to the development of renal disease in SLE. Another morphological relationship between SLE and IFN was reported by S. Rich who found that tubuloreticular structures, that are induced by IFN treatment of human lymphoblastoid cells, are present in both SLE patients and human cancer patients undergoing IFN treatment. J. Hooks (NIH) described the *in vitro* production of IFN- γ by the tumour cells of a patient with an OKT8-positive

Olivine transformed from J. Zussman

J.D. BERNAL (like W.L. Bragg) applied his knowledge of crystal structures to a broad range of topics and, as early as 1936, was able to suggest that the mineral olivine, (Mg,Fe)₂SiO₄, would transform to the spinel structure at the high pressures found deep in the Earth's mantle. There is now much evidence that olivine is the major constituent of the rocks of the Earth's upper mantle, so any transformation that occurs at greater depth is of obvious geological and geophysical importance.

In 1966, Ringwood and Major¹ produced the transformation in the laboratory (at about 1,000°C and 115 kbar) and found not one, but two, high-pressure polymorphs, β - and γ -Mg₂SiO₄. The higher-pressure γ -phase does, as suggested, have the spinel structure and the β -phase a modified version of it. Analogous structures for Co₂SiO₄ were later determined by Morimoto *et al.*². Density increases of about 10 per cent are associated with the transformations and probably explain the discontinuity in seismic P-wave velocity³ observed at a depth of about 400 km. The phase change with increased density may perhaps also be a driving force for the downward movement of slabs of oceanic crust at continental plate margins, and may act as a trigger for deep-focus earthquakes in these regions⁴.

The β - and γ -phases have also been found in certain meteorites and are thought to have been produced by extraterrestrial — rather than Earth impact — shock pressures. Studies of the Tenham chondritic meteorite by Putnis and Price suggested that the β -phase was metastable, but phase relationships in the system Mg₂SiO₄-Fe₂SiO₄ at high pressures and temperatures are particularly complex at the Mg-rich end.

J. Zussman is Professor of Geology in the University of Manchester, M13 9PL.

T-cell leukaemia, and C. Chany has (INSERM) found a variety of IFNs in human amniotic fluids. In addition, L. Epstein discussed several defects in IFN- α or γ production, but saw no reason for treatment of the immunological disturbances in these conditions with IFNs at present, or until there is further understanding of the role of IFNs in the pathogenesis of primary and secondary immunodeficiency diseases.

One purified recombinant DNA-produced IFN, $\alpha 2$, (Hoffman La Roche) stimulated low levels of antibodies in 10 out of 200 treated patients and in this respect, it resembles native IFNs. The National Cancer Institute's Phase I trials of recombinant IFN- $\alpha 2$ and lymphoblastoid IFN also indicate some antitumour activity in various human cancers. In contrast to other published results, an increase in natural killer cell activity could not be

It seems likely that the direct $\alpha \rightarrow \beta$ transformation, and also $\gamma \rightarrow \beta$ and $\gamma \rightarrow \beta \rightarrow \gamma$ can all occur with increasing depth because the associated temperature increase enlarges the stability field of the β -phase⁶.

In this issue of *Nature* (p.729), Price, Putnis and Smith now describe transmission electron microscopic studies of veinlets in the Peace River meteorite. The β -phase occurs in micron-size crystals as pseudomorphs of olivine and also replace ringwoodite (naturally occurring α -Mg₂SiO₄). They identify defects in the ringwoodite as stacking faults which are, in effect, very thin sheets of the β -structure in the fault plane. They point out that the β - and γ -structures can, as a whole, be considered related by such stacking faults, and suggest that growth of the β -phase from the fault plane produces fault-free β -crystals. In the meteorite the β -phase is thus derived from the γ -phase in a post-shock retrograde process occurring on reduction of pressure. This contrasts with the views of Madon and Poirer⁷ that the faults in ringwoodite are associated with the prograde olivine to spinel transformation.

The authors round off with some speculation as to the relevance of the mechanism of transformation that they describe to a $\gamma \rightarrow \beta$ transformation in the Earth's mantle. They suggest that the systems of stacking faults involved might well be sensitive to external shear forces and thus pertinent to the rheological properties of the mantle.

1. Ringwood, A.E. & Major, A. *Earth planet. Sci. Lett.* 1, 241 (1966).
2. Morimoto, N., Tokonami, M., Watanabe, M. & Koto, K. *Am. Miner.* 59, 475 (1974).
3. Ringwood, A.E. *Earth planet. Sci. Lett.* 5, 401 (1969).
4. Ringwood, A.E. *Earth planet. Sci. Lett.* 14, 233 (1972).
5. Putnis, A. & Price, G.D. *Nature* 28, 217 (1979).
6. Akimoto, S., Matsui, Y. & Syono, Y. in *The Physics and Chemistry of Minerals and Rocks*, 327 (Wiley, New York, 1976).
7. Madon, M. & Poirer, J.P. *Science* 207, 66 (1980).

demonstrated. Positive results in cancer patients also come from a French trial of IFN- β in which dosage level and scheduling were the minimum necessary to elevate natural killer activity *in vivo*. Preliminary clinical trials of IFN- β in Germany and Japan were presented, as well as studies of lymphoblastoid IFN produced by the Burroughs-Wellcome group. J. Gutterman (MD Anderson Hospital) reported a series of 19 renal cell carcinoma cases in which 7 patients showed shrinkage of disease with IFN treatment.

Two newly analysed randomized, controlled trials showed positive results. A group from Stanford observed that varicella virus infections in leukaemic children can be improved with IFN and M. Hirsch (Massachusetts General Hospital) reported that the impact of CMV infection on renal transplant recipients can be lessened with IFN- α prophylaxis. □

Breaking the active galaxy speed record

from Andrew Lawrence

OBSERVATIONS of the variability of the light output of quasars and Seyfert galaxies have been of great importance as probes of the physical conditions in these 'active galactic nuclei'. New evidence has been reported¹ that the Seyfert galaxy NGC 6814 varies on the amazingly short time scale of 100 s, not only revealing the smallest dimensions so far for an active galactic nucleus, but at last beginning to place constraints on the physical processes responsible for the prodigious rate of energy output.

The time during which a body can change its behaviour by a significant amount is limited by the speed with which information can travel across the body; large bodies cannot alter as rapidly as small ones. The early realization that some quasars vary on a time scale of months² led astronomers to the astonishing conclusion that bodies that could fit comfortably between here and the nearest star are

emitting as much energy as an entire normal galaxy. The 1970s saw the discovery that quasars and Seyfert galaxies expend a large fraction of their energy budgets as X rays. Several experiments found cases in which X-ray production varied over days and even hours³⁻⁸. Such considerations have led to the common belief that X rays relate most directly to the actual 'central powerhouse' of active galaxies. The volume of space involved in such a fast variation is far smaller than can be directly resolved with either optical telescopes or even the largest baseline radio interferometer. Variability studies may thus extract information unobtainable by any other method.

In 1978, controversy began over evidence presented by a group at Harvard⁹

Andrew Lawrence is at the Royal Greenwich Observatory, Herstmonceux Castle, Hailsham, East Sussex BN27 1RP.

that the X-ray emission from the Seyfert galaxy NGC 4151 had varied on the remarkable time scale of 700 s. The data were rather marginal, and a group at the Goddard Space Flight Center (GSFC) published observations repudiating this suggestion⁵. However, a new study by Tennant *et al.*, from the same group at GSFC¹, has presented convincing evidence that a similar Seyfert galaxy, NGC 6814, varies its output on a time scale of 100 s.

Such very rapid variability tells us more than just the size of the radiating body. It places constraints on the parameters of models suggested to explain the activity of these galactic nuclei. Tennant *et al.* discuss several points in detail.

One hundred seconds corresponds to a light travel distance of 3×10^{12} cm. The X-ray emitter cannot be larger than this. If a massive black hole is at the heart of the business, it cannot be more massive than 10^7 solar masses ($M_{\odot}$); or the event horizon would be outside 3×10^{12} cm. On the other hand, to explain the observed luminosity by accretion of matter requires a black hole of at least $10^5 M_{\odot}$. Such a system can have been radiating at its present luminosity for

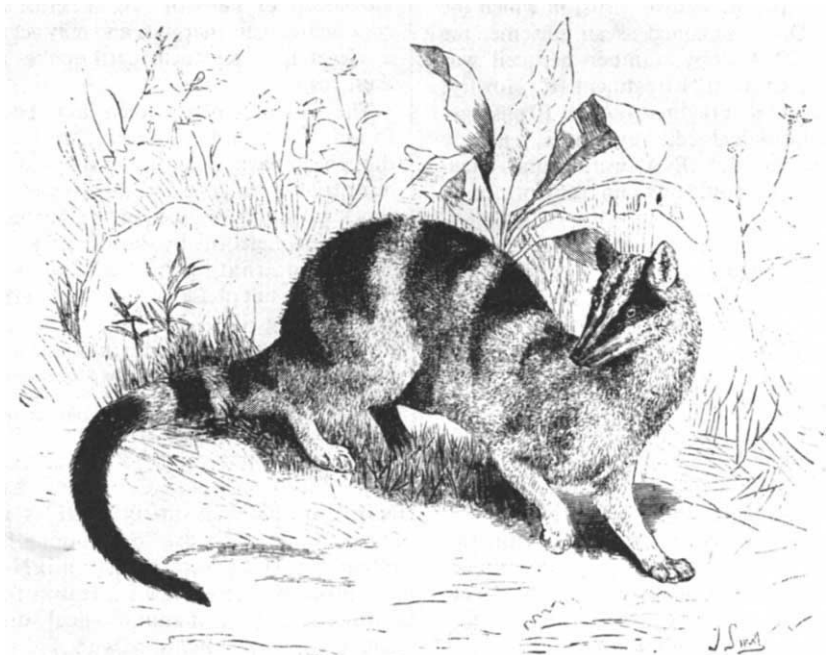


100 years ago

ILLUSTRATIONS OF NEW OR RARE ANIMALS IN THE ZOOLOGICAL SOCIETY'S LIVING COLLECTION

HARDWICKE'S CIVET-CAT (*Hemigalea Hardwickii*). — The Viverridae, or Civet-cats, form a well-marked family of carnivorous mammals peculiar to the tropics of the Old World, and mostly confined to Southern Asia and Africa, though one or two of them occur in the southern parts of Europe. One of the finest and largest of them is the True Civet-cat (*Viverra civetta*), from the anal glands of which the old-fashioned perfume known as civet is extracted, and the Genets, Ichneumons, and Mongooses are well-known members of the same family, examples of which are always to be seen in the Zoological Society's Collection.

Amongst the rarer and less familiar forms of the Viverrine groups is the very curiously-marked animal which we now figure (Fig. 17) from a specimen received by the Society in October, last year. Hardwicke's Civet, though first described by Dr. Gray so long ago as 1830, is a very scarce and little-known species, and the present example is believed to be the first of its kind ever brought alive to this country. In 1840 Müller and Schlegel gave an excellent figure and description of this animal, under the name of *Viverra boiei*, in their great work upon the Natural History of the colonial possessions of the Netherlands. Their specimen was obtained in South-Eastern Borneo by Herr Henrici, and sent alive to the



Gardens of the Zoological Society of Amsterdam. Hence, after its death, it was transferred to the National Museum of Leyden. Müller adds that he never met with this Civet-cat himself during his extensive travels in the Eastern Archipelago, and had received no information as to its habits.

Hardwicke's Civet-cat was also figured and described by Eydoux and Souleyet in the "Zoology" of the voyage of the *Bonite* in 1841, under the name *Hemigate zebra*, but again without any information as to its habits,

not even the locality of their specimen being stated.

So far as has been ascertained from the Zoological Society's living specimen, this animal is excessively shy and retiring in disposition, and apparently does not leave its retreat voluntarily except at night. When handled, it ejects a highly acrid and skunk-like secretion from its anal glands. The length of the body in the example figured is about 24 inches, and that of the tail about 18 inches. From *Nature* 25, 608, 27 April 1882.

only 2 per cent of the currently accepted age of the Universe or it would have accreted more than its own mass. As approximately 2 per cent of all galaxies are active, the new evidence could suggest that all galaxies have an active phase for 2 per cent of their life times.

Much argument about X rays from active galaxies centres on the radiation mechanism involved. Is it thermal? Or is it due to very high-energy relativistic electrons (like those found in radio galaxies)? Such high-energy electrons must lose their energy extremely fast. Tennant *et al.* conclude that for such non-thermal mechanisms to be viable, some process must continuously re-accelerate electrons throughout the emitting volume. Are thermal processes simpler? Not really; the X-ray spectrum is consistent with optically thin thermal emission, but from such a small volume, thermal radiation should be optically thick. Thermal processes are only viable in more sophisticated models where very hot electrons give a 'boost' to UV photons inserted from outside the X-ray-emitting region.

Why have such rapid fireworks not been seen from other objects? In another, as yet unpublished, investigation¹⁰, Tennant *et al.* find that of 30 objects studied, only NGC 6814 showed such behaviour. In other studies of Seyfert galaxies, only NGC 4051 and possibly NGC 3227 have shown variability on a time scale of minutes^{11,12}. All three galaxies are rather low-luminosity examples of the Seyfert phenomenon, but other low-luminosity objects have not shown very rapid variations. There are indications that these objects have anomalous optical and IR properties¹³, but the data are still scant and confusing.

Recently preprinted observations with the Einstein Observatory¹⁴ have claimed 100 s variability in the quasar 1525 + 227. The data are not as convincing as in other cases, but if confirmed, the new observations could be of great importance; the X-ray luminosity of 1525 + 227 is several hundred times that of NGC 6814, and the constraint placed on models would be correspondingly tighter. Hope for further progress in these matters in the near future rests on the performance of the European X-ray satellite EXOSAT, due to be launched in autumn 1982. □

Receptors for T-cell growth

from D.E. Kipp and B.A. Askonas

THE study of lymphocyte mediators has recently become an active area of cellular immunology. Although it has been known for more than twenty years that conditioned media from lectin-stimulated spleen cells contain lymphokines which enhance *in vivo* immune responses^{1,2}, the more recent discovery that conditioned media can support the continuous growth of activated T cells *in vitro*³ has made the T-cell growth factor(s) (TCGF) of particular significance.

In vivo, the importance of TCGF lies in the control of T-cell clonal expansion and responses. *In vitro*, the use of TCGF has proved of enormous value since it has made it possible to select and grow T-cell clones with different functions, such as cytotoxic T cells or T-helper cells. For the first time, single clones of T-cell subpopulations of known antigen specificity can be studied. This has made it possible to characterize their phenotype and functions, and their release of various lymphokines, factors, or interferon; to type lymphocyte-defined human histocompatibility antigens; and to determine the fine specificity of individual T cells and their recognition of antigen in conjunction with products of the major histocompatibility gene complex. It has also made it possible to provide important sources of lymphokines.

The resting T cell is unresponsive to TCGF and presumably does not express receptors for TCGF on its cell surface. In contrast, mitogen- or antigen-activated T-cells respond to TCGF⁴ and continue to grow *in vitro*. In a recent report, R.J. Robb, A. Munck and K.A. Smith⁵ define receptors for TCGF by examining the binding of biosynthetically labelled TCGF derived from a human T-leukaemia cell line (JURKAT) to various cell types. TCGF was purified after incubation of JURKAT cells with radiolabelled amino acids to show a single band on SDS-polyacrylamide gel electrophoresis at a molecular weight of about 15,000, equivalent to the previously estimated molecular size of human TCGF⁶. This was the only measure of purity, but the material was functionally active and its binding properties to cells paralleled functional activity. The authors estimated by equilibrium binding assays of the purified radiolabelled TCGF that there were 10–15,000 receptors on TCGF-dependent mouse cytotoxic or helper T-cell clones. Furthermore, while normal murine splenocytes and thymocytes as well as human peripheral blood cells showed just barely detectable binding of radiolabelled TCGF, alloantigen- or lectin-stimulated cells had

easily detectable receptors. This result supports the observations that only antigen- or mitogen-activated T cells are responsive to TCGF, whereas unstimulated T cells (and lipopolysaccharide-activated B cells) are not reactive to TCGF and appear not to possess any receptors for TCGF.

Robb *et al.* also studied the binding of radiolabelled TCGF to a variety of human neoplastic cell lines. All of the non-T-cell lines and 90 per cent of human T-cell lines failed to bind TCGF. The human T-cell lines have been classed as T-cell lines on the basis of rosette formation with sheep red blood cells and this may not be an ideal criterion. The absence of TCGF receptors may therefore be explained in several ways. The transformed cell lines may have lost expression of their receptors or may have been derived from early T cells that had not yet expressed receptors. Certainly, antigen-dependent untransformed T cells, be they helper or cytotoxic cells, can respond to TCGF^{7,8}. The one human T-cell line found to bind TCGF, which apparently grows well without the addition of exogenous TCGF, has recently been shown to both produce and respond to TCGF⁹.

The work of Robb *et al.* also confirmed the specificity of the human receptor for TCGF. The binding of radiolabelled human TCGF to its receptor was not inhibited by any other growth factor tested or by TCGF derived from rodent spleen cells which does not act on human T cells. On the other hand human TCGF is not species specific and acts on rodent T cells as well¹⁰. The association and dissociation of human TCGF binding to mouse cytotoxic T cells showed a rapid turnover of TCGF and its receptor as well as a rapid degradation of TCGF after association with its receptor. This rapid degradation may have an important regulatory significance as the presence of TCGF would allow for the continued expansion of activated T cells in the lymphoid tissues, and thus block any dampening of the immune response when it is no longer needed.

These studies provide insight into the regulation of the immune response by TCGF. However, it needs to be stressed that this is only the beginning of such studies. Clearly it will be important to define the nature of the receptor for TCGF and we thus await eagerly the next chapter of this unfolding story. □

D.E. Kipp and B.A. Askonas are in the Immunology Division of the MRC National Institute of Medical Research, Mill Hill, London NW7 1AA.

1. Tennant, A. F., Mushotzky, R. F., Boldt, E. A. & Swank, J. H. *Astrophys. J.* **251**, 15 (1982).
2. Burbidge, E. M. A. *Rev. Astr. Astrophys.* **5**, 399 (1967).
3. Elvis, M. *Mon. Not. R. astr. Soc.* **177**, 7p (1976).
4. Winkler, P. F. & White, A. E. *Astrophys. J. Lett.* **199**, L139 (1975).
5. Mushotzky, R. F., Hott, S. S. & Serlemitsos, P. J. *Astrophys. J. Lett.* **225**, L115 (1978).
6. Delvaile, J. P., Epstein, A. & Schnopper, W. H. *Astrophys. J. Lett.* **219**, L81 (1978).
7. Lawrence, A. *Mon. Not. R. astr. Soc.* **192**, 83 (1980).
8. Marshall, N., Warwick, R. S. & Pounds, K. A. *Mon. Not. R. astr. Soc.* **194**, 987 (1981).
9. Tananbaum, H. *et al. Astrophys. J.* **223**, 74 (1978).
10. Tennant, A. F. *et al.* (in preparation).
11. Marshall, F. E., Becker, R. H., Holt, S. S. & Mushotzky, R. F. *Bull. Am. astr. Soc.* **12**, 796 (1980).
12. Mushotzky, R. F. Private Communication.
13. Lawrence, A. & Elvis, E. *Astrophys. J.* (in the press).
14. Matilsky, T., Schrader, C. & Tananbaum, H. Preprint (1982).

1. Gordon, J. & MacLean, L. D. *Nature* **208**, 795 (1965).
2. Kasakuro, S. & Lowenstein, L. *Nature* **208**, 894 (1965).
3. Morgan, D. A., *et al. Science* **193**, 1007 (1976).
4. Smith, K. A., *et al. Molec. Immun.* **17**, 579 (1980).
5. Robb, R. J., *et al. J. exp. Med.* **154**, 1455 (1981).
6. Watson, J. *et al. J. exp. Med.* **150**, 849 (1979).
7. Watson, J. *J. exp. Med.* **150**, 1510 (1979).
8. Gillis, S. & Smith, K. A. *Nature* **268**, 154 (1977).
9. Gootenberg, J. E. *et al. J. exp. Med.* **154**, 1403 (1981).
10. Gillis, S. *et al. J. Immun.* **120**, 2027 (1978).

Shear stress in fault zones

from Barry Kean Atkinson

SINCE estimates of the shear stress acting on fault zones at major plate boundaries vary by at least an order of magnitude, such fundamental issues as the driving mechanism of plate tectonics and the energetics of crustal faulting remain entirely unresolved¹. The shear stresses in the Earth's crust cannot be measured directly, except very close to the surface, so likely stress levels have to be inferred from various secondary lines of evidence.

Laboratory studies of rock friction generally predict that shear stresses on earthquake faults must be at least as high as 100 MPa, unless pore fluid pressures approach that due to the overburden. Such high pore fluid pressures may occur transiently during earthquake faulting through a shear-heating-induced thermal expansion of water² but an impressive array of data (summarized in ref. 3) shows that pore pressures much greater than hydrostatic will generally only develop in crustal rock sequences containing thick blankets of impervious rocks.

A different level of shear stress is indicated by seismological studies. The average stress drop in crustal earthquakes is usually of the order of 10 MPa, or less, irrespective of the strength of the earthquake source, implying that ambient stress levels are also of this order. The key evidence of long-term, low shear stress, probably around 10 MPa, comes from the absence of a heat flow anomaly around the San Andreas Fault. A recent comprehensive review⁴ of the latest heat-flow data for the western US further supports this view. However, unless relative plate velocities are unreasonably high (tens of cm yr⁻¹), the metamorphic gradients and K-Ar ages of rocks along the Alpine Fault of New Zealand are consistent with shear-heating due to a much higher shear stress⁵, of at least 100 MPa.

Geological evidence from fault rock textures suggests that both high shear stress (100 MPa) and low shear stress (10 MPa) faulting occur in the upper crust, with the latter more common⁶. High pore fluid pressures are clearly involved in some low stress crustal faulting and higher shear stresses are usually associated with reverse faulting on dry immature fault zones (the Alpine Fault may be exceptionally dry⁷).

Even if earthquakes along plate boundaries often occur in an environment where ambient shear stress is low, it could be argued that intraplate earthquakes are different and occur where the ambient shear stress is higher. In the eastern US, for

example, some surficial rocks have higher values of deviatoric stress than are found in the western US, perhaps implying shear stresses of several tens of MPa at depth. In addition, the moment of earthquake faulting for a given length of fault break is about one order of magnitude higher in some intraplate regions of the US than in western California. The argument has, however, recently been demolished by Raleigh and Evernden⁸ — for the US at least. They point to some little appreciated facts.

First, attenuation rates of horizontally propagating seismic waves with frequencies relevant to intensity observations (1–4 Hz) are grossly different throughout the US. Second, the energy of intensity-relevant frequencies for earthquakes in the conterminous US is solely dependent on fault length and has no relation to attenuation. Finally, for a fixed length of fault break there is a very substantial increase in seismic moment for a change in attenuation from that typical of western California to that typical of eastern US. Thus, moment values have to be explained within models of the Earth that are highly heterogeneous.

These facts can be explained if earthquakes in eastern US occur along fault zones that constitute soft inclusions in an otherwise rigid and strong crust/mantle system. In order to calculate the true moment of an earthquake, allowance must be made for the effective dimensions of a volume in which stress relaxation occurs. Relatively high moments for short fault breaks are only achieved in association with a large volume of relaxation which can be larger than the inclusion. There is no physical anomaly implied by this result, merely a more heterogeneous earth model than is usually used in seismology.

Raleigh and Evernden⁸ also review the evidence for low shear stress on Californian fault zones and conclude that fault zones in all regions of the US are sites of low ambient shear stress and low stress drop.

Fault motion at low shear stresses has been supposed difficult because high pore fluid pressures have been thought to be the only way to reduce the frictional resistance of a fault. Recent research has shown, however, that the chemical effects of pore water may also have a significant effect.

It is notoriously difficult to run deformation experiments in the laboratory at strain rates comparable with those in tectonically loaded fault zones. Yet, such experiments are essential if the range of potential water-weakening reactions are to be assessed. A few years ago, it was shown that stress relaxation experiments⁹ could be used to assess strain rates on fault zones down to about 10⁻¹⁰s⁻¹, within one or two

orders of magnitude of some tectonic strain rates but some five orders of magnitude slower than typical laboratory strain rates. Since then this technique has been applied to the study of stress levels on water-weakened faults in a range of crustal rocks¹⁰.

For faults in many wet crustal rocks a dramatic weakening occurs at strain rates below 10⁻⁷s⁻¹, probably due to some combination of stress corrosion and diffusional mass transfer. The effect is in addition to any strength reduction due merely to the mechanical effect of high pore fluid pressures. Additional support is given to the stress corrosion hypothesis by the observation that in quartz, the rate of crack growth increases by five orders of magnitude on raising the temperature from 20 to 200°C (ref. 11).

Some of these experimental results can be extrapolated to conditions thought to occur at depths down to 15 km along the San Andreas Fault Zone. For sandstone and quartzite, shear stress for sliding under slow, tectonic strain rates of 10⁻¹¹ to 10⁻¹⁴ s⁻¹ is of the order of 10 MPa, even if pore water pressure never exceeds hydrostatic. For granite, however, a modest pore water overpressure is required to lower shear stress for sliding into the 10 MPa range.

The explanation for diversity of shear stresses on crustal fault zones may lie in water's different mechanical, thermodynamic and, especially, chemical properties. There is sufficient variation both in the time scale of fault zone processes and in the material properties of crustal rocks which, given the diverse physico-chemical actions of water, could explain the range of shear stress estimates for fault zones. There is a strong case⁸ that the environment in which many earthquakes occur is of low shear stress, but with stronger asperities of varying sizes and density of distribution along the fault surface which account for less common, local shear stress of several tens of MPa.

Future work must concentrate on working out the details of potential chemical weakening effects of pore water on rock strength under simulated crustal conditions. Some reliable means will also be required for extrapolating laboratory experiments to the larger strains typical of slip on earthquake fault zones. □

Barry Kean Atkinson is in the Department of Geology, Royal School of Mines, Imperial College, Prince Consort Road, London SW7 2BP.

1. Hanks, T.C. & Raleigh, C.B. *J. geophys. Res.* **85**, 6083 (1980).
2. Sibson, R.H. *Nature* **243**, 66 (1973).
3. Brace, W.F. *Int. J. Rock Mech. Min. Sci. & Geomech. Abstr.* **17**, 241 (1980).
4. Lachenbruch, A.H. & Sass, J.H. *J. geophys. Res.* **85**, 6186 (1980).
5. Scholz, C.H., Beavan, J. & Hanks, T.C. *J. geophys. Res.* **84**, 6770 (1979).
6. Sibson, R.H. *J. geophys. Res.* **85**, 6239 (1980).
7. Sibson, R.H., White, S.H. & Atkinson, B.K. *Bull. R. Soc. N.Z.* **18**, 55 (1979).
8. Raleigh, C.B. & Evernden, J. *Am. geophys. Un. Geophys. Monogr.* **24**, 173 (1981).
9. Rutter, E.H. & Mainprice, D.H. *Pageophysics* **116**, 634 (1978).
10. Atkinson, B.K. & Dennis, S.M. *Trans. Am. geophys. Un.* **62**, 1036 (1981).
11. Meredith, P.G. & Atkinson, B.K. *Trans. Am. geophys. Un.* **62**, 1040 (1981).

REVIEW ARTICLE

Atomic and gravitational clocks

V. M. Canuto*[†] & I. Goldman*

* NASA Goddard Institute for Space Studies, Goddard Space Flight Center, New York, New York 10025-7886, USA

† Department of Physics, City College of New York, New York 11031, USA

Atomic and gravitational clocks are governed by the laws of electrodynamics and gravity respectively. While the strong equivalence principle (SEP) assumes that the two clocks have been synchronous at all times, recent planetary data seem to suggest a possible violation of the SEP. Our past analysis of the implications of an SEP violation on different physical phenomena revealed no disagreement. However, these studies assumed that the two different clocks can be consistently constructed within the framework. The concept of scale invariance, and the physical meaning of different systems of units, are now reviewed and the construction of two clocks that do not remain synchronous—whose rates are related by a non-constant function β_a —is demonstrated. The cosmological character of β_a is also discussed.

OF the forces of nature, the two more successfully described are the electromagnetic forces, through quantum electrodynamics (QED) and macroscopic gravity, through Einstein's theory of general relativity (GR). The high level of agreement between predictions and observations leave little doubt that we now possess the correct physical interpretation as well as the theoretical tools to describe both forces.

QED and GR are also complete theories in the sense that they yield operationally well-defined clocks which satisfy the dynamical equations of the theories themselves. To understand what is required for a theory to yield a well defined clock, we introduce the notion of scale invariance, SI.

Consider a dynamical equation defined in a given system of units, containing field variables and dimensional parameters. Consider now a scale (length) transformation of the type

$$L \rightarrow L_* = \Omega_*^{-1}(x)L \quad (1)$$

where $\Omega_*(x)$ is a dimensionless arbitrary function of space-time. Regarding coordinates as dimensionless space-time markers, it follows from equation (1) that

$$g_{\mu\nu}^* = g_{\mu\nu} \Omega_*^{-2}(x), \quad ds_* = \Omega_*^{-1} ds \quad (2)$$

Furthermore, a physical tensor Λ of arbitrary rank will be taken to transform like

$$\Lambda \rightarrow \Lambda^* = \Lambda \Omega_*^{-\pi(\Lambda)} \quad (3)$$

where $\pi(\Lambda)$ is called the power of Λ : $\pi(ds) = 1$, $\pi(g_{\mu\nu}) = 2$. Because null cones transform into null cones, it follows that $c = 1$ holds in all systems of units. Therefore, because $v = (v/c)c$, $\pi(v) = 0$. The power of any quantity can therefore be expressed in terms of $\pi(m)$ and $\pi(L) = 1$.

Given these definitions, an equation is scale invariant if: (1) in the transformed system of units, it preserves the same form involving the transformed fields; (2) it has the same parameters; and (3) there is no explicit dependence on $\Omega_*(x)$.

We now show that a SI theory cannot yield a unit of time, that is, a clock. Because of the assumed SI, the dynamical equations governing the clock are identical in any system of units and so their solution for the period p of the clock in one given system of units is the same in any other system of units, $p = p_*$. On the other hand, if instead of solving the dynamical equations, we apply the transformation (1) directly to p , we obtain the result $p = \Omega_* p_*$, that is $p \neq p_*$, in contradiction with the previous result. It follows that a SI theory yields a solution

that is simultaneously constant and variable, thus proving that such a quantity is physically meaningless. In other words, a SI theory, being invariant with respect to changes in scale, does not possess a scale, thus lacking an intrinsic unit of time.

Therefore, a SI theory cannot provide a clock. For a clock to exist, the underlying theory must be non SI.

As an example, let us consider the case of the electromagnetic clock, an electron revolving around a proton, governed by the following equations¹:

$$\frac{1}{\sqrt{g}} [\sqrt{g} F^{\mu\nu}]_{,\nu} = J^\mu, \quad u^\mu{}_{;\nu} u^\nu = \frac{e}{m} u^\lambda F^\mu{}_\lambda \quad (4)$$

The periodic solution with the period p is given by

$$p = 2\pi \frac{m^2 l^3}{e^4} [1 + O(v/c)^2] \quad (5)$$

where l is the angular momentum per unit mass. In equations (4) there is a 'time coordinate' x^0 whose physical meaning is not given by equations (4), where x^0 is in fact a marker. Its physical meaning can be determined by considering that since $de/dx^0 = dm/dx^0 = dl/dx^0 = 0$, we have from equation (5) $dp/dx^0 = 0$, thus allowing one to attribute a physical meaning to x^0 . Equations (4) are therefore constructed so as to yield a physical unit of time, given by equation (5). This property is due to the fact that equations (4) are not SI. In fact, let us apply equation (1) to equations (4). As $2\pi(e) = \pi(h) = 1 + \pi(m) = 2 - \pi(G) \equiv 2 - g$ and $2\pi(F^\mu{}_\nu) = 2\pi(E, H) = \pi(\rho_\gamma) = \pi(m) - 3$, (E and H are the electric and magnetic field strengths, ρ_γ is the energy density, e^2/hc is a pure number and both h/mc and GM/c^2 have the dimensions of a length), equations (4) become treating coordinates as dimensionless,

$$\frac{1}{\sqrt{g_*}} [\sqrt{g_*} F_*^{\mu\nu} \Omega_*^{4+\pi}]_{,\nu} = J_*^\mu \Omega_*^{4+\pi} \quad (6a)$$

$$u_*^\mu{}_{;\nu} u_*^\nu + \Delta_*^{\mu\nu} \frac{\Omega_*^{\pi\nu}}{\Omega_*} = \frac{e}{m} u_*^\lambda F_*^\mu{}_\lambda \Omega_*^{\pi+3} \quad (6b)$$

with $\Delta_{\mu\nu} \equiv u_\mu u_\nu - g_{\mu\nu}$ and $2\pi \equiv 2\pi(F^{\mu\nu}) = \pi(m) - 7$.

Clearly, while equation (6a) can be made SI by assuming $\pi(m) = -1$, equation (6b) cannot, for the Ω_* dependence in the Δ term cannot be made to disappear because there are no free parameters in $\Delta_{\mu\nu}$. Therefore the system (4) is not SI.

Let us now consider a gravitational clock (a planet and a star, for example) and let us consider the Einstein equations describing macroscopic gravity. From the previous analysis it follows that the lack of SI of equations (4) is due solely to the equations of motion which in the gravitational case are already contained in the Einstein field equations, which can therefore be expected to be not SI. This is indeed the case²: Einstein equations do not retain their original form under the scale transformations (1)–(2). Their lack of SI is why a gravitational clock is a well-defined quantity, whose period, given by Kepler's third law, reads (M is the total mass and J is the orbital angular momentum per unit mass)

$$P = \frac{2\pi J^3}{(GM)^2} \quad (7)$$

The above discussion shows that the electromagnetic and gravitational clocks are meaningful because the dynamical equations governing them are not SI.

The strong equivalence principle

To study the relationship of the two clocks, consider a physical phenomenon characterized by a proper length interval Δs , which we shall measure using electromagnetic and gravitational clocks, the results being Δs_a (atomic) and Δs_E (Einstein), respectively.

Because clocks are the manifestation of underlying forces and we do not yet possess a unified theory, we do not know *a priori* whether the ratio

$$\frac{\Delta s_E}{\Delta s_a} = \beta_a \quad (8)$$

is constant. The lack of knowledge of the function β_a has so far been circumvented by adopting the strong equivalence principle^{3,4} (SEP) which assumes that

$$\beta_a = \text{constant} = 1 \quad (9)$$

implying that, for example, the value of the period of a planet should be independent of the clock used to measure it.

The SEP comprises two assumptions³: (1) that the weak equivalence principle (WEP) holds; and (2) that local gravitational and non-gravitational experiments are independent of where and when in the Universe they are performed. The requirement (9) refers to 'when in the Universe', since, due to the high degree of homogeneity observed in the Universe, we assume that β_a is only time dependent. Furthermore one experiment is not sufficient to test the SEP, at least two experiments are needed at two different times, because β_a can be normalized to unity at any one time; what is physically relevant is β_a .

Tentative evidence of a violation of SEP

Using 8,249 lunar occultation measurements⁵, lunar radar ranging data^{6,7}, and dynamical determinations of the lunar periods^{8–10}, it has been suggested that at the present epoch $\beta_a/\beta_a \approx 10^{-11} \text{ yr}^{-1}$. However, as tidal forces make the Earth–Moon system less than an ideal laboratory for our purposes, it would be more satisfactory to use radar ranging data to the inner planets Mercury, Venus and Mars. Using the available results, an upper limit $|\beta_a/\beta_a| \leq 10^{-10} \text{ yr}^{-1}$ was set¹¹.

We stress here one aspect of the theoretical analysis. It may seem natural to use the standard Einstein equations with the simple alteration $G \rightarrow G(t) = G_0 + \dot{G}_0 t + \dots$, whenever G appears. However, from Einstein equations it follows quite generally that for an isolated system $GM = \text{constant}$ (refs 12, 13), where M is the total mass. (For any local system, this constraint holds regardless of cosmological expansion³.) The violation of this constraint can lead to serious errors. For example, because the period and distance to a planet are given by $P \sim J^3/(GM)^2$, $R \sim J^2/(GM)$, $GM = \text{constant}$ implies $\dot{R} = 0$ and $\dot{P} = 0$ and not $2\dot{R}/R = \dot{P}/P = -2\dot{G}/G$, as usually stated. This

argument shows that a value of $\dot{G}/G$ cannot be extracted from the standard Einstein equations where by construction G always appears multiplied by M , the product being required to satisfy the constraint $GM = \text{constant}$.

For these reasons, we (in collaboration with R. Helling and P. J. Adams of JPL) have enlarged the system of dynamical equations (originally used by R. Helling) to make them compatible with a possible violation of the SEP. The Viking radar ranging data to Mars are now being analysed. The best fit solution will hopefully provide a reliable value for β_a .

Purpose and basic assumptions

Let us construct a theoretical framework that allows for an SEP violation in the form of a non-constant β_a . This allows for the presence of two fundamental systems of units in nature, which is necessary to explore consistently the implications of an SEP violation as well as to provide relations which can be subjected to observational test.

Gravitational or dynamical units (EU, where E stands for Einstein), are the units in which Einstein equations remain unchanged. These field equations contain the equations of motion that yield the clock equation (7), which is therefore taken to give the gravitational or dynamical unit of time (also called ephemeris time). In EU, $G_E = \text{constant}$ by definition and so the total mass is also constant, since $G_E M_E$ is constant. In addition, from $T_{E;\nu}^{\mu\nu} = 0$ applied to a pressureless fluid, we find that the rest mass is also constant. Therefore, in general

$$G_E = \text{constant}, \quad M_E = \text{constant} \quad (10)$$

where M is either total or (macroscopic) rest masses. As the particle number N will be shown not to be constant, equation (10) does not mean that microscopic masses m_E are constant in EU. Other atomic quantities such as e and h , are also in principle not constant in EU.

Atomic units (subscript a) are the units in which the dynamical behaviour of an electromagnetic clock is governed by equation (4), which in turn implies that equation (5) is taken to be the atomic unit of time. Furthermore, in AU we have

$$e_a = \text{constant}, \quad m_a = \text{constant}, \quad h_a = \text{constant} \quad (11)$$

The first two terms are the analogue of (10) for microphysics. (Actually, terms in equation (11) are contained in equations (4).) In analogy with what we said earlier, the second term of equation (11) does not imply that macroscopic masses M_a are constant in AU. In fact, they are not (see equation (25)); G_a is also not constant (see equation (18)).

Let us now consider the basic problem of determining the relation of the two 'preferred' systems of units. We introduce a language that describes any physical equation in a general system of units of which the two 'preferred' systems are special cases.

Physics in general units

Let us reconsider equation (1) and transform L_* to L_{**} ,

$$L_* \rightarrow L_{**} = \Omega^{-1}(x) L_* = (\Omega_* \Omega)^{-1} L \equiv \Omega_*^{-1} L$$

The quantity Ω_* has therefore power -1 , as from equation (3), where Λ^* , Λ and Ω_* are replaced by Ω_{**} , Ω_* and Ω , respectively. Therefore, any quantity of the form $\Lambda^* \Omega_*^{\pi(\Lambda)}$ has zero power under subsequent scale transformations:

$$\Lambda^* \Omega_*^{\pi(\Lambda)} = \Lambda^{**} \Omega_{**}^{\pi(\Lambda)}$$

As an example, we perform a further transformation of equation (6) to a g_{**} , F_{**} , J_{**} system. The final result can easily be seen to be of the same form as equation (6), with only double starred quantities in it. Therefore, equation (6) can be said to be written in general units.

Let us now define a fiducial system of units by $\Omega_* = 1$, and a general system of units by $\Omega_* = \beta$. Here, as in previous work¹⁴,

we choose EU as the fiducial system ($\beta_{\text{EU}} = 1$), so that to be consistent with equation (8), the AU system is defined by $\beta_{\text{AU}} = \beta_a$. Note that while β , defining a general system of units, has power -1 , β_a is of power zero, because $\beta_a = \beta_{\text{AU}}/\beta_{\text{EU}}$ is the same in all systems of units. Clearly the use of general units does not introduce any new physics.

The action

To deal with the problem of constructing the two clocks, we propose an action in general units, which as such must be of zero power,

$$I = \int \mu \beta^{2-g} ds + \frac{1}{16\pi} \int \beta^{2-g} F_{\alpha\gamma} F^{\alpha\gamma} \sqrt{g} dx^4 + \int \beta^{2-g} e A_\nu u^\nu ds \quad (12)$$

where μ represents masses in general (microscopic or/and macroscopic) and where the relation between $F_{\mu\nu} \beta^{1-g/2}$ and $A_\alpha \beta^{1-g/2}$ is the usual one. It is easy to check that the power of I is zero. Because the dimensions of I are $[M][L]$, the β^{2-g} factors are required for $\pi(I) = 0$. The matter part of equation (12) is different from that proposed by Dirac¹⁵ which is a particular case of equation (12) if $\mu \beta^{1-g} = \text{constant}$. In AU, and for microscopic masses, this implies $g = 1$, because of equation (11). However, $g = 1$ will be shown not to allow the two clocks to run at different rates. The relaxing of the restriction $\mu \beta^{1-g} = \text{constant}$ is why we can construct two clocks that run at different rates.

Macroscopic gravitational clock

Consider the periodic motion of a planet in the gravitational field of a star. From the first term in equation (12), we derive the following equations of motion in general units,

$$u^\alpha_{;\gamma} u^\gamma + \frac{(\mu \beta^{2-g})_{,\gamma}}{(\mu \beta^{2-g})} \Delta^{\alpha\gamma} = 0 \quad (13)$$

where the metric $g_{\mu\nu}$ due to the star is given in general units by the Schwarzschild metric times β^{-2} . As we are dealing with a macroscopic object, then

$$\mu = M, \quad M_E = \beta^{1-g} M = \text{constant} \quad (14)$$

where the second relation is the general law for mass transformation from Einstein units to general units, following from equation (3) with $\Lambda = M_E$, $\Lambda^* = M$, $\Omega_* = \beta$, $\pi(m) = 1 - g$. The last equality in equation (14) follows from equation (10). Equation (13), with equation (14), specializes to

$$\text{EU: } u^\alpha_{;\gamma} u^\gamma = 0; \quad \text{AU: } u^\alpha_{;\gamma} u^\gamma + \frac{\beta_{a,\gamma}}{\beta_a} \Delta^{\alpha\gamma} = 0 \quad (15)$$

The solutions for the period P can be easily worked out. The results are

$$\begin{aligned} \text{EU: } P_E &= \frac{2\pi J_E^3}{(G_E M_E)^2} = \text{constant}; \\ \text{AU: } P_a &= \beta_a^{-1} P_E \end{aligned} \quad (16)$$

Electromagnetic clock

The motion of an electron in the field of a proton $F^{\lambda\nu}$, is governed by the following two equations, derivable from equation (12)

$$(\sqrt{g} F^{\lambda\nu} \beta^{1-g/2})_{,\nu} = 4\pi \int e \beta^{1-g/2} \delta^4(x^\alpha - z^\alpha) dz^\lambda \equiv J^\lambda \quad (17a)$$

$$u^\alpha_{;\gamma} u^\gamma + \frac{(\mu \beta^{2-g})_{,\gamma}}{(\mu \beta^{2-g})} \Delta^{\alpha\gamma} = \frac{e}{\mu} u^\nu F^\alpha_\nu \quad (17b)$$

In equation (17b) we used $e \beta^{1-g/2} = \text{constant}$, a constraint derivable from equation (17a) using the antisymmetry of $F_{\mu\nu}$. Let

us now specify equation (17) to AU and EU. In AU we require equation (17) to coincide with equation (4). In AU, $\beta = \beta_a$, and for a microscopic mass $\mu = m = m_a = \text{constant}$. The requirement can therefore be fulfilled only if

$$g = 2, \quad \pi(m) = -1, \quad G_a \beta_a^2 = \text{constant} \quad (18)$$

To derive the last term of equation (18) we have used equation (3) with $\Lambda \equiv G_E$, $\Lambda^* \equiv G_a$, $\Omega_* \equiv \beta_a$, as well as equation (10).

The corresponding solution for the period p_a in a local lorentzian coordinate frame is now given by equation (5), with the subscript a attached to all the quantities. Let us now consider EU, where

$$\begin{aligned} \beta &= 1, \quad \mu = m_E = m_a \beta_a^{\pi(m)}, \\ e^2 &= e_E^2 = e_a^2 \beta_a^{1+\pi(m)} \end{aligned} \quad (19)$$

where we have again used equation (3). Because $g = 2$, equation (17a) retains the same form as in AU, whereas equation (17b) becomes

$$\text{EU: } u^\alpha_{;\gamma} u^\gamma - \frac{\beta_{a,\gamma}}{\beta_a} \Delta^{\alpha\gamma} = \frac{e_a}{m_a} \beta_a u^\nu F^\alpha_\nu \quad (20)$$

Using a local lorentzian coordinate frame, the solution for the period p_E is found to be

$$p_E = \beta_a p_a, \quad (p_a = \text{constant, equation 5}) \quad (21)$$

To achieve the desired result we had to fix a gauge: a relation between β_a and G_a , equation (18). This is a welcome feature because until now, we had to consider g as a free parameter^{16,17}. This no longer the case, as the theory now demands equation (18). Note that such a gauge was previously suggested in connection with the 3K black-body radiation¹⁸.

We have proposed a lagrangian whose solution for the periods of gravitational (P) and atomic (p) clocks are

$$P_E = \beta_a P_a, \quad p_E = \beta_a p_a, \quad P_E, p_a = \text{constant} \quad (22)$$

or

$$\frac{p_a}{P_a} = \frac{p_E}{P_E} \sim \beta_a \quad (23)$$

namely: in either atomic or gravitational units, the ratio of the periods of the two clocks is not constant, provided β_a is not constant. We have therefore proved that, provided $g = 2$, a framework exists which allows the two clocks to run at different rates.

We must stress that the extension of equation (17a) to a charged fluid must be written as $F^{\mu\nu}_{;\nu} = \tilde{J}^\mu$, where $\tilde{J}^\mu = \epsilon u^\mu f$, f being an undetermined function of β_a . Due to the antisymmetry of $F^{\mu\nu}$, it follows that $\tilde{J}^\mu_{;\mu} = 0$, which implies $\epsilon N f = \text{constant}$. Because for $g = 2$, e is constant in any units, equation (19); it then follows that $f \sim N^{-1}$. The Coulomb force now becomes $e^2 N^2 f^2 / r^2$. The analogous gravitational force is GM^2/r^2 , with $GM^2 = G_E M_E^2 = \text{constant}$, a result valid for $g = 2$. The correspondence between macroscopic Coulomb's and Newton's laws is therefore preserved.

Weak equivalence principle

In achieving the result (23), the central part was played by the action (12) and the equations of motion ensuing from it. Since equation (13) is in general units, using $g = 2$, the fact that $m \beta^{1-g} = m_E = m_a \beta_a^{1-g} \sim \beta_a^{-1}$, and equation (14), we obtain

$$\text{Microscopic bodies: } u^\alpha_{;\gamma} u^\gamma + \frac{\beta_{a,\gamma}}{\beta_a} \Delta^{\alpha\gamma} = \frac{\beta_{a,\gamma}}{\beta_a} \Delta^{\alpha\gamma} \quad (24a)$$

$$\text{Macroscopic bodies: } u^\alpha_{;\gamma} u^\gamma + \frac{\beta_{a,\gamma}}{\beta_a} \Delta^{\alpha\gamma} = 0 \quad (24b)$$

which indicate that the two types of bodies do not follow the

same type of trajectories. Equations (24) are still in general units.

From the operational point of view, we are only interested in AU, so we limit our considerations to them. Equations (24) tell us that in AU, microscopic masses follow geodesics while macroscopic masses do not. In either case, however, masses do not enter the equations and the WEP is separately satisfied, in the sense that all macroscopic objects follow the same path as do all microscopic ones. The results of the Eotvos–Dicke–Braginskii experiments^{3,19}, showing that two (macroscopic) bars of Al and Au (or Al and Pt) follow the same path, are therefore in full agreement with equations (24), as the extra 'force' represented by β_a , is independent of the mass and it affects both bars equally. Thus, its effect cancels out in this type of experiment.

To test equations (24) one should follow in time the trajectories of an atom and of a planet, which is the procedure in the radar ranging experiments. In fact, one may think of atomic and gravitational clocks (the Earth–Moon system), as two 'objects' moving in space–time following two given trajectories. If, as time evolves, the two systems follow different types of trajectories, charting the time evolution of the macroscopic object with the reference provided by the microscopic object cannot yield constant results if the ratio of the proper lengths spanned by the two objects is not constant in time. Therefore equations (24) are an alternative way of interpreting the lunar and planetary data that stresses the difference in the two trajectories rather than the difference of the two clocks. These two ways of interpreting the data are equivalent, although the second one is the one almost exclusively referred to in this context.

Particle number non-conservation

The fact that only $g = 2$ is allowed has important consequences. In fact, since m_a and M_E are constant, it follows that (using equation (3) between AU and EU)

$$m_E \sim \beta_a^{1-g}, \quad M_a \sim \beta_a^{g-1}, \quad N \sim \beta_a^{g-1} \quad (25)$$

implying that N is no longer constant. To have a conserved N , we have to choose either $\beta_a = \text{constant}$, in which case the SEP is automatically satisfied or $g = 1$, which would also lead us to an SEP conserving framework. In fact, for $g = 1$ the left-hand side of equation (15) governing the macroscopic gravitational clock in AU would be identical to that of equation (17b) governing a microscopic electromagnetic clock in AU, thus leading to no difference between the two periods P_a and p_a , thus returning to the SEP. (The right-hand side of (17b) does not affect this statement because, due to spherical symmetry, it does not affect the angular momentum conservation law).

Equation (25) is the most important consequence of the SEP violation framework as it indicates that a violation of the SEP is intimately related to a violation of the particle number conservation law.

Gravitational constant

It is often stated that if atomic and gravitational clocks are different, the gravitational constant G must vary with atomic time. While the statement is not incorrect, it may give the impression that it adds some new fact. This is not the case. In fact, neither in the gravitational action of ref. (14) nor in the one presented here, is there an independent function G . One calls G the combination $G_E \beta^{-g} = G$, $G_E = \text{constant}$. But one does not introduce new physics, one simply lumps together a function β , a parameter g , and a constant G_E . That the physics is contained in β_a and not in G_a , is evident from the fact that the experiments on the Moon and the inner planets, yield directly β_a and not $\dot{G}_a$, which is derived quantity, equation (18).

Cosmological meaning of β_a

The framework presented here while permitting the examination of the implications of an SEP violation, does not explain the physical mechanism behind it. In fact, β_a is treated here as an external quantity, much as viscosity is treated in classical fluid mechanics, where it enters as an external parameter whose evaluation requires a microscopic kinetic theory.

Although we do not offer a dynamics for β_a , it is important to stress that a dynamics of β_a can be either of a local or global nature. In the first case, β_a is regarded as a space–time field described by an action to be added to the total action; this would entail a coupling of β_a to local matter, with the result that even in EU, macroscopic gravity would no longer be described by standard Einstein equations, thus departing completely from our basic assumptions. A local approach was adopted by Brans and Dicke (BD). As several studies have indicated^{20,21}, Solar System experiments constrain, within the BD approach, the variability of β_a to some orders of magnitude below the value quoted previously. As there is no reason why we should arbitrarily restrict β_a to such low values, we find the local approach inadequate.

The value of $\dot{\beta}_a$ implying $\dot{\beta}_a/\beta_a \sim H_0$ suggests that β_a is related to the structure of the Universe and that its dynamics is likely to be governed by topological rather than local space–time considerations. The SEP violation represents therefore a cosmological influence on local physics, in accord with Mach's principle. In contrast, accepting the SEP as an exact law of physics is equivalent to assuming that local physics is independent of the rest of the Universe.

Recent work on nucleosynthesis²² has indicated that, as expected, an SEP violation with a time scale of the order of the Hubble time cannot be extrapolated back to the radiation dominated era. Furthermore a dynamics for photons can be constructed²² independently of β_a : in particular, a very general argument has been found indicating that the photon number N_γ , contrary to the particle number N , equation (25), is adiabatically conserved for any value of the parameter g . (The photon treatment presented in (ref. 16) is therefore valid only if $g = 1$).

The two previous results suggest that an SEP violation, if it exists, began to manifest itself only after the Universe entered the matter dominated era, before which β_a may have been constant.

Conclusions

The most important results of the present analysis are:

- Einstein field equations retain their standard form only in EU. In AU, they depend on β_a and their form is given by equation (2.34) of ref. 14.
- The trajectory of a macroscopic (many body) object is a geodesic in EU; in AU, β_a factors enter. The results are given in equation (24b).
- The trajectory of a microscopic (one body) object is a geodesic in AU; in EU, β_a factors enter, equation (24a).
- While in AU the description of a microscopic one body dynamics is unchanged, (this holds true even at the level of the Schrödinger equation), the description of a many-body situation is affected by β_a . In fact, the particle number $N \sim \beta_a$, equation (25). This in turn implies that macroscopic masses are such that $M_a \sim \beta_a$, $M_E \sim \text{constant}$. Microscopic masses are such that $m_a \sim \text{constant}$, $m_E \sim \beta_a^{-1}$. Finally, the gravitational coupling G is such that $G_E = \text{constant}$, $G_a \sim \beta_a^{-2}$.

We thank Drs P. J. Adams, J. Anderson and J. Lodenquai for constructive criticism, and Ms Doris Smith for typing the manuscript.

1. Weinberg, S. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972).
2. Singe, J. L. *Relativity: The General Theory* Chap. VIII (North-Holland, New York, 1976).
3. Will, C. M. in *General Relativity* (eds Hawking, S. W. & Israel, W.) (Cambridge University Press, New York, 1979).
4. Thorne, K. S., Lee, D. L. & Lightman, A. P. *Phys. Rev. D* **7**, 3563 (1973).
5. Van Flandern, T. C. *Astrophys. J.* **248**, 813 (1981).

6. Dickey, J. O., Williams, J. G. & Yoder, C. F. *EOS* **61**, 46, 939 (1980).
7. Calame, D. & Mulholland, J. D. *Science* **199**, 977 (1978).
8. Morrison, L. V. & Ward, C. G. *Mon. Not. R. astr. Soc.* **173**, 183 (1975).
9. Goad, C. C. & Douglas, B. C. *J. geophys. Res.* **83**, 2306 (1978).
10. Lambeck, K. *Phil. Trans. R. Soc.* **287**, 545 (1977).
11. Reasenberg, R. D. & Shapiro, I. I. in *On the Measurements of Cosmological Variations of G* (ed. Halpern, L.) (University Presses of Florida, Gainesville, 1978).
12. Landau, L. D. & Lifshitz, E. M. *The Classical Theory of Fields* 2nd edn (Addison-Wesley, New York, 1962).
13. Tolman, R. C. *Relativity, Thermodynamics and Cosmology* (Oxford University Press, New York, 1966).
14. Canuto, V. M., Adams, P. J., Hsieh, S.-H. & Tsiang, E. *Phys. Rev. D* **16**, 1643 (1977).
15. Dirac, P. A. M. *Proc. R. Soc. A* **333**, 403 (1973).
16. Canuto, V. M. & Hsieh, S.-H. *Astrophys. J. Suppl.* **41**, 243 (1979).
17. Canuto, V. M., Hsieh, S.-H. & Owen, J. R. *Astrophys. J. Suppl.* **41**, 263 (1979).
18. Canuto, V. M. & Hsieh, S.-H. *Astrophys. J.* **224**, 302 (1978).
19. Rudenko, V. N. *Soviet Phys. Usp.* **21**, 11, 893 (1978).
20. Weinberg, S. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* 628–629 (Wiley, New York, 1972).
21. Bekenstein, J. *Comments Astrophys.* **8**, 89 (1979).
22. Canuto, V. M. & Goldman, I. Preprint (NASA Goddard Institute for Space Studies, New York, 1982).

ARTICLES

Protein dynamics investigated by the neutron diffraction–hydrogen exchange technique

A. A. Kossiakoff

Department of Biology, Brookhaven National Laboratory, Upton, New York 11973, USA

A new approach, using neutron diffraction and the hydrogen exchange (H/D) technique, has been used to study the extent and nature of the inherent conformational fluctuations in the protein, trypsin. The observed pattern of exchange was used to investigate systematic relationships between exchangeable sites and structural and chemical properties of the molecule. Results of this analysis indicate that hydrogen-bonding structure is the dominant factor governing rates of exchange. The model of conformational mobility which best explains the experimental findings involves a localized disruption of the secondary structure within different regions of the protein molecule, each limited in extent to the breaking of a small number of hydrogen bonds.

STRUCTURAL and chemical investigations have clearly demonstrated that, despite their high degree of hydrogen and hydrophobic bonding, proteins are far from rigid matrices of atoms and have certain component segments which exhibit considerable dynamic mobility^{1–5}. Because of their fundamental importance to the understanding of biological processes^{6–8}, the dynamical properties of proteins have recently been the subject of intense study by a variety of methods. However, important issues remain concerning the effects of protein folding on the extent and nature of its dynamical motions.

I report here a study of the mobility characteristics of trypsin, a 223 amino acid protein, by coupling the hydrogen exchange (H/D) technique with neutron diffraction. Since its introduction by Linderstrom-Lang and his colleagues^{9,10}, the H/D exchange method has been widely recognized to have great potential as a probe for protein conformational change (for detailed discussions of the potential applications of the H/D method to evaluating protein dynamics see refs 2, 11–14). H/D exchange has major advantages over other labelling techniques in that a deuteron has a negligible space requirement and has chemical properties nearly equivalent to those of the proton it replaces in the structure. Further, potentially labile sites are distributed throughout the molecule and therefore probe the dynamical properties of the whole structure.

Kinetic studies have shown that H to D exchange rates of peptide NH groups in a protein can differ by as much as 10 decades¹². It has been generally supposed that the groups which exchange at rapid rates similar to those characteristic of small molecules are those located on the surface of the protein and hence in direct contact with the solvent. Conversely, the groups which exchange much more slowly are assumed to be located in the interior of the protein and involved in internal hydrogen bonding^{11,12}.

It is generally accepted that in the conditions of this experiment (pH 7), the exchange reaction is catalysed by the presence of hydroxyl ion. However, little is known about other relation-

ships between the exchange chemistry and the dynamical properties of proteins. For instance, the size of the solvent unit required to facilitate the exchange process is an open question, a situation which has led to the postulation of two distinct stereochemical models for exchange. One model, which will be referred to as the cooperative or local unfolding mechanism, is described by a transient, cooperative unfolding of a segment of secondary structure^{11–13}. Access to the exchangeable protons is assumed to be accomplished by the extrusion of the chain into the bulk solvent where the exchange reaction can proceed by normal water chemistry. In the second model, the penetration mechanism, it is assumed that the exchange reaction takes place, shielded completely from the bulk solvent, within the tightly packed interior of the protein. It is proposed that the necessary solvent molecules can be diffused through the protein to the interior sites via pathways opened by local atomic fluctuations¹⁴ or by mobile defects in the protein packing¹⁵.

A variation of local unfolding is introduced here and termed 'regional melting'. It differs from the mechanism described above in that exchange does not require the chain to be extruded into the bulk solvent. In this variant, the reformation process is limited to the cooperative breaking of several hydrogen bonds, resulting in the formation of a solvent-filled cleft in the protein surface. The exchange is assumed to take place within the cleft if the cleft is sufficiently large to permit the solvent molecules to arrange themselves in a stereochemically productive orientation with respect to the exchange site. Although the solvent molecules within the clefts cannot be considered to have chemical properties identical to the bulk phase, they are assumed to be contiguous to the bulk solvent.

Unfortunately, previous H/D exchange experiments have been unable to relate exchange rates with specific groups or even regions of polypeptide chain, thus limiting the usefulness of exchange methods in elucidating the factors responsible for protein conformational mobility, and the mechanism of the exchange reaction.

Recently, the application of NMR spectroscopy¹⁶⁻¹⁸ and chemical analysis^{13,19,20} to H/D exchange has begun to provide data relating exchange rates to structural features. The present article reports on the application of neutron diffraction to this problem. Two factors make this technique particularly suitable for providing a direct and quantitative approach to the measurement of H/D exchange rates of specific sites. First, because it is a crystallographic technique, the precise location of each labile site in the well-ordered segments of the protein is known and can be examined directly in a neutron Fourier map. Second, because the amplitudes of hydrogen and deuterium scattering are of opposite signs (Fig. 1), the process of assigning a labile site as having either H or D character is quite a straightforward task, provided that the structure of the protein has been highly refined, as is the case for trypsin.

As will be seen from subsequent discussion, the goal of this study is not simply to obtain dynamical information about the trypsin molecule *per se*, but to use trypsin's structural units, its β -sheet and α -helix regions, its turns and loops, to investigate through H/D exchange how each type of unit reacts to the conformational mobility modes of the protein. Correlations between the exchangeability of a potentially labile site and four principal factors are examined: (1) the degree of local hydrogen bonding structure around the site, (2) its distance to the solvent, (3) the hydrophobic or hydrophilic character of adjacent side chains and (4) its observed vibrational motion (temperature factor).

Crystalline proteins as models for dynamic studies

It could be objected that, since diffraction methods are limited to the study of crystalline proteins, they might be expected to yield different results from studies of proteins in solution. The use of a crystalline system to model a dynamic phenomenon like conformational mobility is valid because a protein molecule in a crystal is highly solvated and has an environment very similar to that in solution²¹⁻²³. Thus, crystallographic studies have shown that a protein crystal can best be described as an ordered and open array of molecules held together by a relatively small number of intermolecular contact points. The interstitial region between molecules generally makes up about one-half of the total crystal volume²⁴, and is comprised of a network of continuous channels filled with solvent having properties analogous to bulk water^{23,25,26}. Because of the extent of unoccupied volume in the crystal, ions and substrates can readily diffuse through the solvent channels to interact with the protein. Analysis of the extent of the solvent network in the trypsin crystal ($\approx 46\%$ solvent by volume) revealed several channels having cavities as large as 18 Å across.

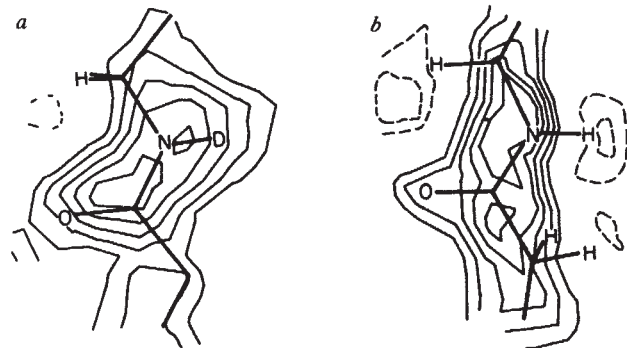


Fig. 1 Sections of a neutron density map taken in the plane of the peptide group. *a*, Fully exchanged peptide group, scattering length of D = $+6.7 \times 10^{-13}$ cm; *b*, unexchanged peptide group, H = -3.8×10^{-13} cm. In cases where partial exchange has taken place, the expected density at the H/D site is close to zero. This is because with H and D having scattering amplitudes of opposite sign, the composite scattering density tends to cancel in the Fourier summation. In the refinement, H/D sites were assigned a temperature factor based on that of the amide nitrogen, and occupancy factors alone were refined. In those regions of the structure with a high degree of thermal motion or local disorder, H/D ratios were not considered reliable and the affected residues are identified in Fig. 2 legend.

Given this structure, the protein molecules themselves would be expected to exhibit many of the chemical traits found in their fully solvated state. Particularly relevant to the present study are several investigations comparing the H/D exchange kinetics between crystalline and solvated protein systems²⁷⁻³⁰. The exchange properties of solution and crystal systems were found to be very similar, implying that accurate exchange information can indeed be obtained from a crystalline protein system.

To determine whether or not lattice packing forces at contact areas between adjacent molecules in the crystal produced a measurable degree of inhibition of the exchange reaction, the peptides involved in the principal contact zones were identified and their degrees of exchange examined. All but one of the sites were found to be completely exchanged (the lack of exchange at this site is probably due to its being in a β -sheet structure and not a result of the contact), which demonstrated that such inhibition as might be encountered in the molecular contact regions was not significant for the reaction time of this experiment. However, this analysis does not exclude the possibility that some of the more extensive modes of motion might in some ways be impeded by the lattice forces.

Experimental procedures

A 5-mm³ crystal of monoisopropylphosphoryl (MIP) trypsin was soaked in a D₂O solution at pH 7 and 20°C for ~ 1 yr before data collection was begun. (The MIP group is a small covalently bound inhibitor which does not affect the tertiary structure of the protein³¹.) Data to 1.8 Å resolution (14,500 reflections $> 3\sigma$) were collected on a two-dimensional position-sensitive detector system³² at the Brookhaven High Flux Beam Reactor. These data were processed by a method developed in this laboratory to deal specifically with the problems inherent in obtaining accurate intensity measurements from the weak neutron diffraction of protein crystals³³. The initial phasing model was calculated by applying the appropriate neutron scattering lengths to the refined trypsin X-ray coordinates of Chambers and Stroud³⁴. Refinement of the structure was carried out by the restrained difference Fourier technique, which involves the determination of proper atomic shifts by the curvature-gradient method and then the re-idealization of the structural parameters by an energy minimization routine. Details of these experimental procedures are reported elsewhere³¹. The current trypsin model is highly restrained to ideal bond lengths and angles (r.m.s. error in bond lengths < 0.014 Å, bond angles $< 1.4^\circ$). After eight cycles of refinement the *R* factor is 19.2%.

The peaks produced in a neutron density map by sites having intermediate exchange ratios (between 15% D–85% H and 60% D–40% H) were found to give densities of a magnitude comparable with the general noise features of the trypsin map; therefore, prediction of relative exchange ratios between sites within this range cannot be made with statistical confidence. Accordingly, it is appropriate to subdivide the exchange ratios into three major categories: (1) unexchanged (0–15% D), (2) partially exchanged (15–60% D) and (3) fully exchanged (60–100% D). Limiting the number of categories in this manner will retain the informational content of the H/D data while guarding against over-interpretation of the relevance of small changes in exchange ratio.

Patterns of exchange

The degree of exchange of the amide hydrogens, as determined from the neutron diffraction analysis, is represented schematically in Fig. 2. Three-dimensional views of the main-chain atoms of the molecule represented in space filling and ball and stick form are shown in Fig. 3, which illustrates the exchange pattern as a function of the structure and folding of the protein. Side-chain groups were omitted to permit an unencumbered view of the intramolecular main-chain hydrogen bonding.

In summary, of the 215 exchangeable amide groups, 68% were found to be fully exchanged, 8% partially exchanged and 24% unexchanged. This corresponds to the cumulative result

of the exchange process throughout the entire span of the D_2O soaking of the crystal, because once a D replaces an H in the structure, any further exchange at the site is between Ds because of the low concentration of hydrogens in the soaking solution.

Although this experiment only distinguishes sites which exchange fast or slowly relative to the fixed time of reaction (1 yr at pH 7, 20°C), the relatively low fraction of partially exchanged sites indicates a very substantial difference between the degree of protection provided by structural features of the protein to the unexchanged sites and that afforded the exchanged sites.

Inspection of Fig. 2 reveals that the unexchanged sites form a distinctive pattern whose most prominent feature is the clustering of unexchanged sites in regions corresponding to the β -sheet structures of the protein. The interior core of trypsin consists of two such structures, called ' β -barrels', each consisting of six antiparallel strands laced together by a network of hydrogen bonds. In fact, 45 of 52 unexchanged hydrogens are found in the β -sheet structures. Closer examination of these regions shows that only 11 of the peptide groups involved in β -sheet hydrogen bonding are completely exchanged, and all of these are located at edges of sheet structures. Modelling studies performed by Salemme³⁵ show that extended β -sheet structures can withstand substantial deformation without disrupting their hydrogen bonding network. The present findings are consistent with this observation. Note, however, that the structural integrity of a β -sheet unit is not wholly a hydrogen-bonding phenomenon, but also involves hydrophobic forces.

Significant data regarding the accessibility of sites within the tightly bonded β -sheet core come from the exchange behaviour of the hydroxyl group of the Ser 54 side chain. As illustrated in Fig. 4, which shows a segment of buried β -sheet (residues 43–45, 53–56), the Ser OH group forms hydrogen bonds to 43 O and 55 NH. Thus, it is embedded in the main β -sheet network and consequently has the same degree of protection as its neighbouring peptide groups. This hydroxyl group is found to be almost fully exchanged, while the amide proton of residue 55 is unexchanged, as are its neighbours.

The exchange of the OH group, which is considerably more reactive than peptide NH groups^{11,12}, demonstrates that conformational fluctuations large enough to expose the β -sheet region to attack by the solvent do occur, but with insufficient probability for the slower peptide exchange to progress appreciably within the time frame of the experiment.

In addition to considering structural and bonding effects, it is important to examine and discuss several other parameters which might be expected to affect the exchange process, notably solvent accessibility, the chemical nature of side chains, and observed vibrational motion.

Effect of solvent proximity on rates of exchange

The degree of exchange of sites located near the surface of the protein molecule would generally be expected to be greater than that of sites buried in the interior. To test the magnitude of this effect, the distance of each peptide proton to the nearest bulk solvent interface was computed. Figure 5a plots this distance against the degree of exchange. The average distance between peptide sites and the protein surface was found to be 3.8 Å.

As might be anticipated, >90% of unexchanged sites were located 4 Å or more from the surface. However, as many as 27% of the fully exchanged sites were also located 4 Å or deeper. Because access to such sites would presumably require rather substantial conformational fluctuations, the 20 exchanged sites located 5 Å or deeper were examined to determine whether they were near one of the small clusters of water bound in cavities inside the molecule. Nine of these sites were, in fact, located adjacent to one of several interior water molecules. Thus, if the distance to the solvent interface is redefined to include interior water cavities, the average distance

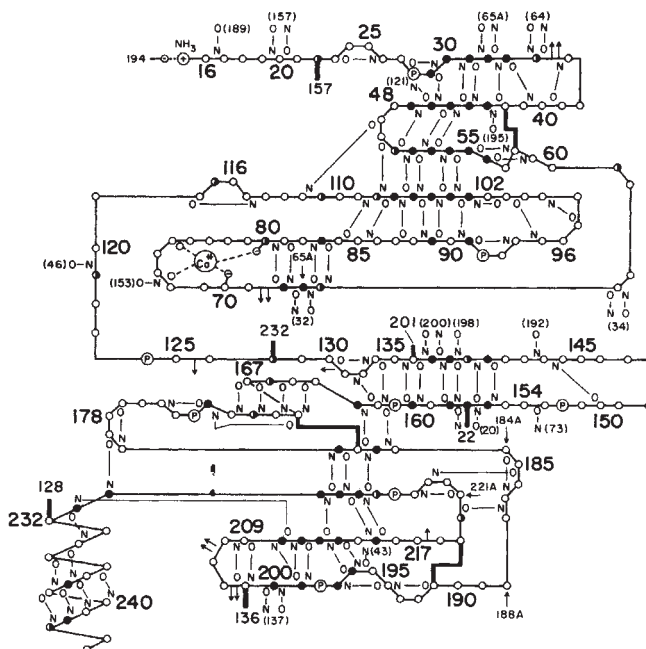


Fig. 2 Schematic representation of H/D exchange at each amide peptide site in the trypsin molecule. Key: ○, full exchange; ◐, partial exchange; ●, unexchanged; P, proline; ↓, sequence insertions; ↑, deletions (based on the chymotrypsinogen numbering scheme); ⊖, carboxylate side chains. Peptide NH and carbonyl O atoms are shown when H-bonded. Disulphide bridges are indicated by broad lines. Because of the effects of thermal motion or local disorder, the H/D exchange information for residues 60–62, 110–111 and 147–150 is not reliable. (Figure design adapted from Birktoft and Blow⁴⁶.)

of these sites from solvent decreases by 2.65 Å (only two sites remain as far as 5 Å from solvent) and the overall average distance of all peptide protons in the molecules from solvent decreases by 1.1 Å to 2.7 Å.

Recent investigations have shown that trypsin is not a unique protein in having a significant number of interior water cavities (I. D. Kuntz, personal communication). The finding that the water molecules residing in the cavities are fully exchanged for D_2O suggests that free solvent exchange occurs between the bulk and interior solvent regions. Figure 6, which shows a low-resolution (5 Å) electron density model of trypsin, gives some idea of the extent of these cavities. Note that, because of the limited resolution of the data and the density contour level used in the construction of the model, the actual sizes of the cavities are exaggerated. Even so, the 5 Å density model still presents a very instructive picture regarding the extent of the cavity network in the trypsin molecule. (The entrances to the pockets holding the interior water molecules are marked in Fig. 6.) Further, a majority of the holes pictured in the 5 Å model were found to coincide with the location of the set of cavities determined from the refined 1.8 Å structure.

In several cases these water cavities are sealed off from direct contact with bulk solvent by only one or several side-chain interactions; the reorientation of the side chains to accommodate interior-bulk solvent contact would presumably be of relatively low energy compared with structural fluctuations which break main-chain hydrogen bonds. Hence, it is plausible that structural fluctuations of the groups surrounding these internal cavities might represent natural low-energy pathways for solvent access and 'regional melting'.

It was noted above that a strong correlation exists between non-exchanged peptides and hydrogen-bonding sites in the β -sheet structures. As the β -structures are located deeper within the protein (3.8 Å) than the remaining sections (2.3 Å), an effort was made to separate out the effects of structure and depth. The average depth of the 11 fully exchanged sites located in the β -sheet structures is 4.0 Å, which corresponds closely

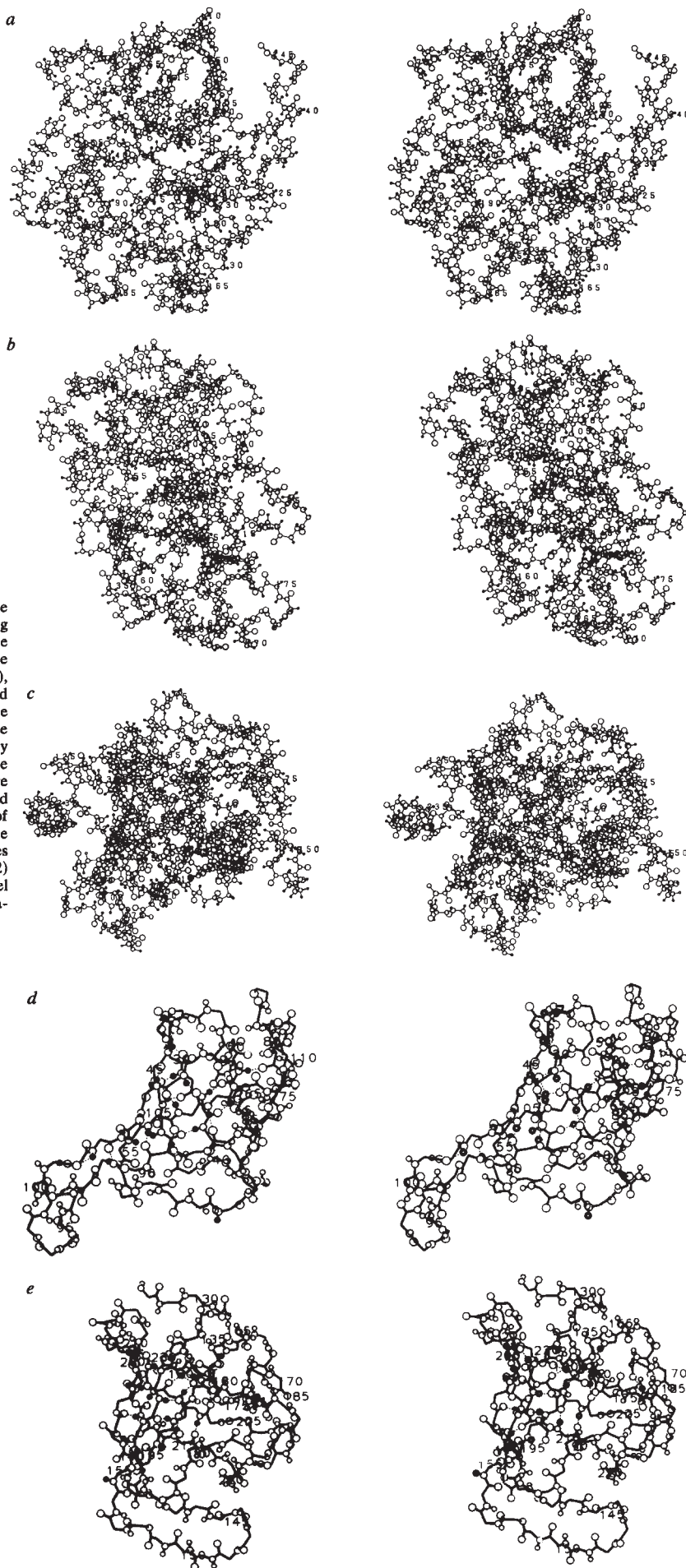
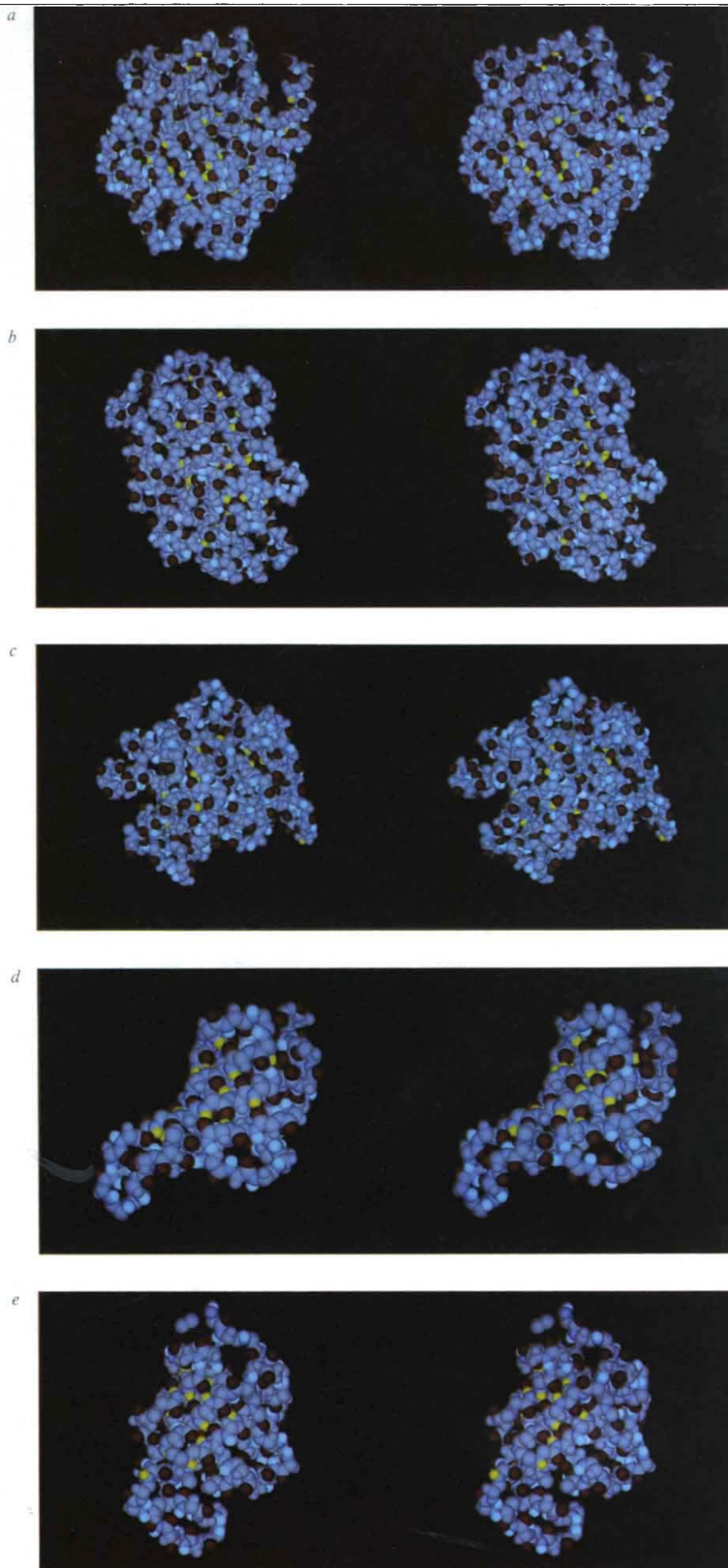


Fig. 3 Stereo space-filling drawings showing locations of the exchangeable amide peptide sites in trypsin. The corresponding views with ball and stick representation are shown on the opposite page. Only main-chain atoms are included (and the C_{α} H atoms). The colour coding is S, N, C, C_{α} , H_{α} (charcoal), carbonyl oxygen (red), unexchanged sites (yellow), exchanged sites (light blue) and partially exchanged sites (white). *a*, Whole molecule, the C-terminal α -helix extends up the right-hand side of the molecule. *b*, Whole molecule (side view), looking directly at the C-terminal α -helix. *c*, Whole molecule (top view), the helix is on the left side of the molecule. *d*, β -barrel structure (1) composed of residues 27–113. The β -barrel is composed of the six strands schematically shown in the upper half of Fig. 2. In the ball and stick representation unexchanged amide protons are designated by heavy-lined circles, with dashed lines denoting their H-bonding interactions. *e*, β -barrel structure (2) composed of residues 129–230. The six strands of the barrel are shown in the lower half of Fig. 2. Ball and stick representation same as in *d*.



to the average depth (3.8 Å) of the total β -sheet population and implies that distance to the solvent is not a primary factor in determining exchange characteristics of β -structures.

In the non- β -sheet sites, the 133 fully exchanged sites were located at an average distance of 2.1 Å from the solvent, while the 7 unexchanged sites were at 4.3 Å. This suggests that the depth effect may be a real one, but the number of unexchanged sites is too small to provide confidence in the result. Note that even after the interior water cavities are taken into account, many sites 3 and 4 Å from the solvent interface are fully exchanged. This would imply that the molecule has sufficient internal mobility to provide access to sites when they are not immobilized by secondary structures.

A graphic example of the dominant role of secondary structure on exchange characteristics is illustrated in Fig. 7, which shows a region of β -sheet containing an occluded water molecule. In the structure, a pair of hydrogen bonds is formed between peptides 215 and 227; the amide hydrogens are unexchanged in both peptide groups. There is a well-ordered water molecule perpendicular to the plane of the sheet which hydrogen bonds to both carbonyl oxygens and to Ser 190 O_γ (not shown). The clear implication is that an amide peptide group can participate with water (in this case D₂O) in a hydrogen bond to the same acceptor ligand and yet remain unexchanged. (This is not an isolated case, since similar arrangements have been observed in five places in the trypsin molecule.) In the above instances there is clearly solvent access to peptide sites which nevertheless remain unexchanged, indicating that structural (bonding) effects dominate over the effect of solvent accessibility.

Neighbouring side-chain effects

It has been suggested that the different chemical characteristics of side chains might, by inductive effects³⁶ or by changing the

local water chemistry³⁷⁻³⁹ around the neighbouring peptide site, significantly affect the rate of H/D exchange. The importance of inductive effects was verified by Molday *et al.*³⁶ who systematically evaluated the nearest-neighbour side-chain effects using small model peptides and derived a set of empirical rules which could predict rates of exchange in random order peptides from their amino acid sequence. Most of the hydrophobic and hydrophilic side-chain effects were found to span a 10-fold range. Other studies have shown that exchange in model polyamides is slowed by the presence of bulky apolar groups, even though the peptide nitrogen itself was not involved in hydrogen bonding³⁹⁻⁴¹. In the light of these investigations, it is important to ascertain whether or not patterns of exchange in the trypsin molecule are also affected by the nature of the associated side chains.

The observed character of these near-neighbour effects³⁶ might suggest that the presence of a hydrophobic side chain would make the adjacent nitrogen more resistant to exchange. In trypsin, the groups alanine, leucine, isoleucine, methionine, phenylalanine and valine account for 65 of the 215 potentially exchangeable sites. Of these, 44% were found to be fully exchanged, 8% partially exchanged and 48% unexchanged. Of the hydrophilic groups—arginine, asparagine, aspartic acid, glutamine, glutamic acid, histidine and lysine—which account for 53 sites, 79% were exchanged, 8% partially exchanged and 13% unexchanged.

Comparison of above values with the average for the entire molecule (68% fully exchanged, 8% partially exchanged and 24% unexchanged) indicates a strong correlation between the degree of exchange and the polarity of the associated side chain. However, before concluding that this correlation is significant, it is necessary to determine whether this effect is independent of the structural correlation previously observed between β -sheet structures and exchange character.

There are 59 sites involved in the two six-stranded β -sheet structures. Of these, 29 are hydrophobic, while only 11 are hydrophilic. (Such hydrophobic/hydrophilic ratios are typical of β -structures of most globular proteins⁴² and it is generally assumed that the high concentration of hydrophobic groups contributes to the unusual stability of these structures.) These β -sheet sites account for all but seven unexchanged groups. Thus, when the members of the β -sheet structures are removed from the hydrophobic population, a large percentage of the remaining groups are fully exchanged (72% fully exchanged, 8% partially exchanged and only 20% unexchanged). Conversely, of the 11 hydrophilic groups which participate in β -sheet structures, 27% are fully exchanged, 9% partially exchanged and 64% unexchanged.

The above analysis suggests that in the conditions of the experiment (1 yr soaking), any effect of side-chain character which might either promote or inhibit exchange is dominated by the more powerful effects of structure and the hydrogen-bonding character of the sites.

Temperature factor correlations

Investigators have attempted to gain insight into the phenomenon of dynamic mobility by using temperature factor data from detailed X-ray analyses of several proteins^{3,4,43}. The attraction of this method is that it identifies specific groups in the protein structure whose locations show positional variation due to thermal motion. However, the method has several limitations. Protein dynamics as described by temperature factor correlations involve small atomic displacements (of the order of 0.5–1.5 Å), and thus represent as a class rather low-energy conformational fluctuations. Further, at the resolution level of normal protein structure analysis (that is, below atomic resolution), accurate individual temperature factor values (*B* factors) are difficult to obtain. Local trends of group *B* factors offer a considerably more reliable measure of motion and, therefore, during the refinement procedure, *B* factors are usually restrained in some fashion to be similar to those of their near neighbours.

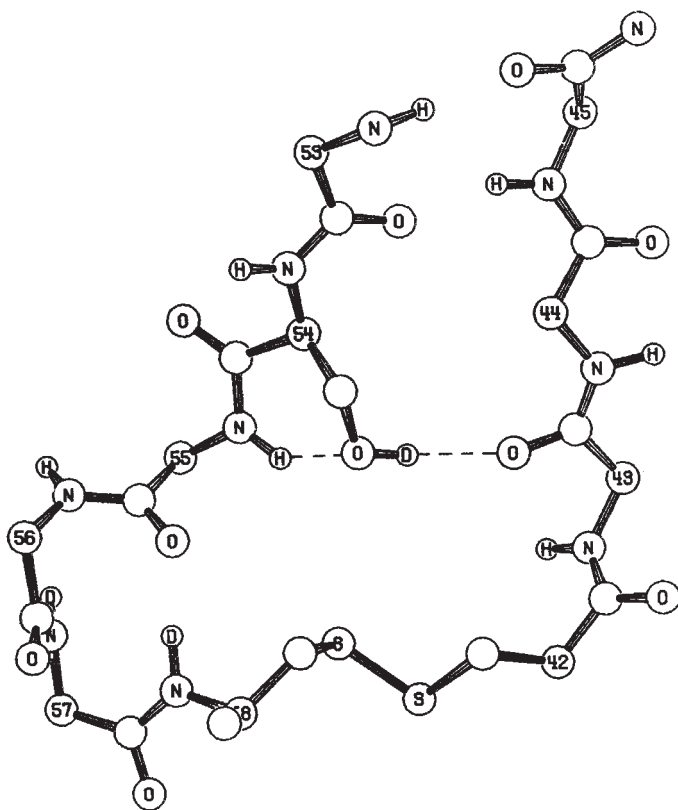


Fig. 4 Section of β -sheet structure residues 43–45 and 53–56 showing main-chain atoms (and the side chain of Ser 54). Because of a bulge in the sheet, the H-bond interaction between 43 O and 55 NH cannot be made directly. The bonding structure of the sheet is maintained (dashed lines) through the mediation of the side chain of Ser 54. The hydroxyl proton of the serine is almost fully exchanged for a D, yet the amide proton of Ala 55 to which the hydroxyl oxygen is H-bonded remains unexchanged.

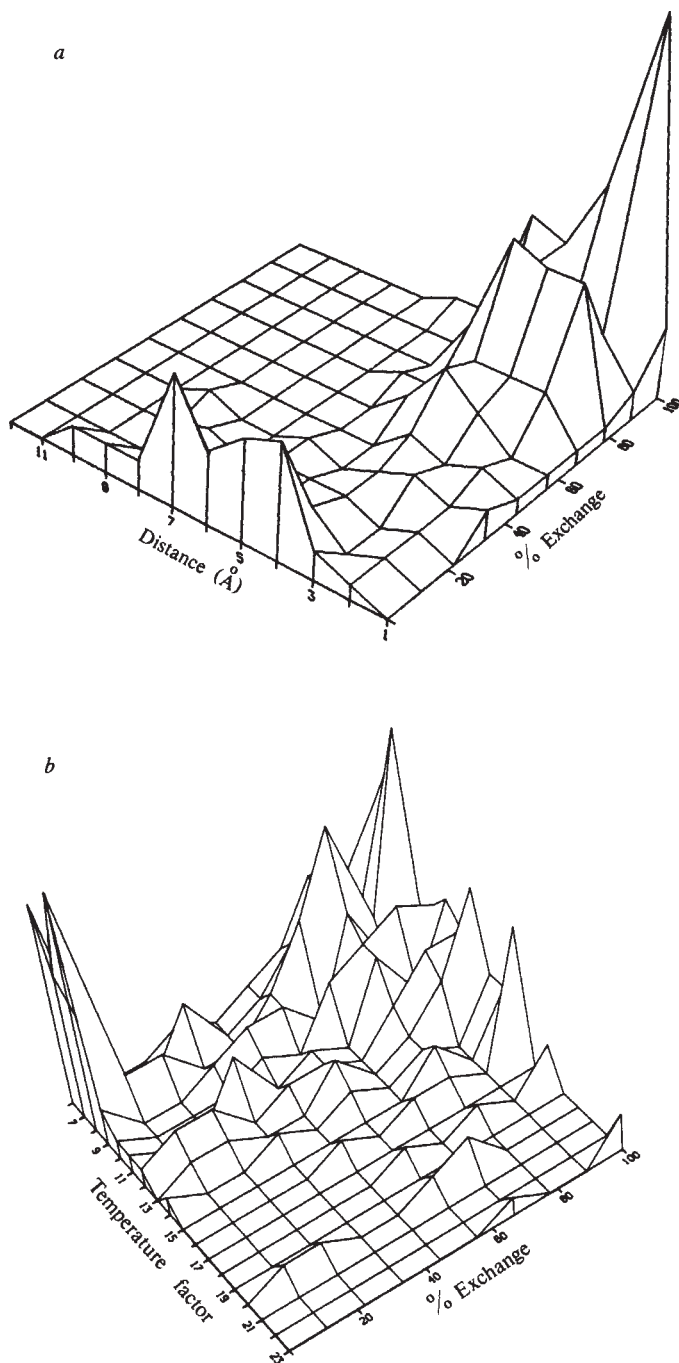


Fig. 5 *a*, Exchange ratios (0–100%D) of the peptide NH sites plotted as a function of their distance to the bulk solvent. Internal water cavities were not considered in the analysis. The molecular envelope was calculated using an atomic radius of 1.4 Å for all atom types including hydrogen. A radius of this size, when applied to hydrogen, overestimates the spatial volumes of most groups and therefore tends to extend the envelope outside the bounds of the molecule defined by the coordinates of the static structure. However, it was felt that in this type of analysis it was important to enlarge the molecular envelope slightly to take into account the small vibrational motions and positional disorder of the surface residues, effects which also increase slightly the molecular boundary. Thus, the resulting site to solvent distances are defined here to reflect a maximum distance, that is, a distance of 1 Å represents a true distance of 0–1 Å to solvent; a distance of 2 Å is 1–2 Å and so forth. Note that in most instances solvent proximity, as defined here (shortest distance to the solvent interface), would not be expected to correlate with solvent accessibility exactly because the low energy pathway by which solvent gains access to a site does not necessarily correspond to the same pathway which defines the shortest distance. *b*, Plot of exchange ratios of the peptide NH sites as a function of their experimentally determined temperature factors. Exchange categories are partitioned as follows: (1) unexchanged (0–15% D), (2) partially exchanged (15–60% D), (3) fully exchanged (60–100% D). Heights of the peaks correspond to the number of sites at each occupancy grouping. Highest peak in *a* (which is distance 2 Å and 100% D) represents 25 sites, in *b* a temperature factor of 8 at 0% D represents 10 sites.

Sternberg *et al.*³ attempted to use thermal factor parameters to describe correlated rigid body motilities of extensive segments of polypeptide chain, applying a sophisticated mathematical treatment to assess the modes of intramolecular motion (so called hinge-bending motions) between the two major globular lobes in lysozyme. From their analysis they conclude that temperature factor information offers only weak evidence for any hinge-bending motions. This suggests that if the coupled modes of movement between the two lobes exist in lysozyme they are structural fluctuations of the high-energy, short-lived type.

The conclusions of Sternberg *et al.*³ are corroborated by the findings of the present neutron analysis. Figure 5*b* shows the correlation between individual amide peptide temperature factors and H/D exchange ratios. As expected, the unexchanged amide groups are distributed in a narrow range at low temperature factors, indicating that they are predominantly located in interior, highly structured regions of the protein. It is somewhat surprising, perhaps, to find that many of the fully exchanged sites are also distributed near the low end of the temperature factor scale. This suggests that they, too, are highly ordered, at least on a crystallographic time scale, but must become sufficiently exposed to solvent as a function of the general breathing modes of the protein to promote exchange.

Because many exchanged interior sites have *B*-factor values comparable with those of the unexchanged sites, I conclude that no distinct correlation exists between the potential exchange of an interior site and its observed thermal motion. This finding has an important implication with respect to the use of crystallographically determined temperature factors to describe protein motions. Although these parameters have been shown to be of value in identifying local dynamic trends involving small atomic displacements, they appear to contain no systematic information on the characteristics of the larger-scale breathing modes of this protein.

Concluding remarks

It has been reported that the variation in the rate of H/D exchange of peptide groups in a protein molecule ranges over as much as 10 decades^{1,2}. The conditions of the experiment reported here were apparently such as to exchange all except those sites which were especially well protected by the protein structure. Clearly, an important future study would involve a comparative analysis of a crystal soaked for a shorter time to identify those sites which exchange slowly on an absolute scale, but rapidly enough to be fully exchanged in the conditions used here.

In summary, it appears that essentially all the sites in which the peptide hydrogens are bonded directly to water molecules—either in the bulk solvent regions or in interior clusters—are fully exchanged. No significant inhibition of peptide exchange was apparent at molecular contact points in the crystal, at hydrophobic side chains or as a function of degree of vibrational motion of the site, except when any of these also corresponded to hydrogen-bonded structures. Further, most of the sites in which the peptide hydrogens are involved in isolated bonds to either side-chain groups or main-chain carbonyls are fully or partially exchanged. Finally, those sites in the β -sheet structures which are not only hydrogen bonded to main-chain carbonyls, but are also flanked by similarly bound peptides, are to a very large extent unexchanged.

In the two short segments of α -helix in the trypsin molecule, the peptides are almost wholly exchanged. The only fully unexchanged sites, Ile 238 and Ile 242, are located along the interface between the helix and two strands of β -sheet. The spatial relationship between these groups and the sheet can be seen in Fig. 3*a*. (As pictured, the helix is located on the right side of the molecule with the amide protons of 238 and 242 shown in yellow.) The finding that the exchange properties of the helix peptide nitrogens differ substantially (apparently depending on whether or not hydrophobic intramolecular contacts are made) indicates that such conformational fluctuations

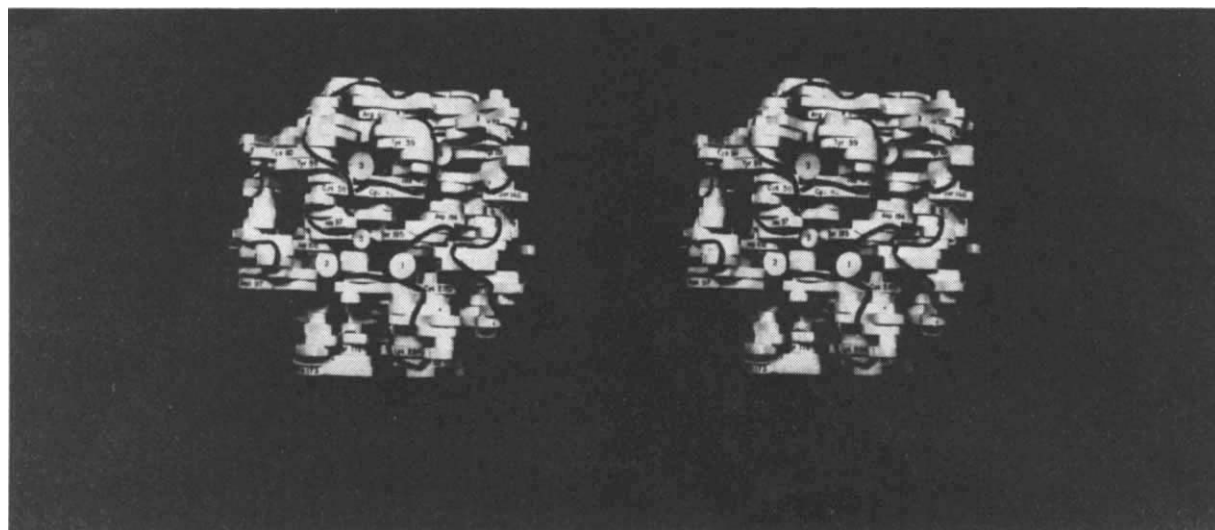


Fig. 6 5 Å density model of trypsin. Water cavity pockets are numbered. Cavity number 1 is the entrance to the substrate binding pocket.

as may be responsible for exposing the groups to solvent probably involve the breaking of hydrogen bonds only adjacent to the exchanged sites rather than the cooperative unfolding of the full helix unit. If such an unfolding process were to occur, it would be difficult to explain how these two particular sites remain unexchanged (especially in light of the fact that they both fit a distinctive pattern with respect to their packing against the β -sheet).

Figure 7 illustrated the important point that solvent accessibility alone is not a sufficient condition for exchange. The finding that rates of exchange are highly correlated to the structural features of the protein, especially hydrogen bonding, is consistent with the hypothesis that the segments of a protein molecule continually undergo local conformational fluctuations in which interior sites occasionally become aligned in chemically reactive orientations with solvent molecules. This mechanism predicts that the relative exchange rate constants would be dependent on the equilibrium concentrations of the unfolded species, which in turn would be a function of the energy required to overcome the local internal bonding of the folded structure. However, there is no evidence that extensive reformation of the polypeptide chain is necessary to expose even the well buried exchanged sites. The finding that the exchangeable sites are partitioned almost exclusively into just two categories (fully exchanged and fully unexchanged, with relatively few sites being partially exchanged), implies that significantly different energies are required to overcome the internal bonding within the various domains in the molecule. Similar findings have been obtained through other experimental techniques^{44,45}.

The fact that full exchange occurs at regions of direct contact of adjacent molecules in the crystal, where gross motion is certainly minimized, indicates that outside the β -sheet structures relatively limited displacements of the structure suffice to provide the necessary access to enable exchange to proceed. This is also borne out by the complete exchange of internally occluded water molecules. Although it seems likely that solvent cavities provide natural pathways for access to some of the deeply buried sites in trypsin, the principal results of this experiment are inconsistent with a mechanism based on the premise that depth of burial of a site is a primary factor governing its rate of exchange. The most likely model to explain the observed exchange behaviour seems to be one involving localized conformational mobility, where the most probable fluctuations are those characterized by a degree of 'regional melting' usually limited in extent to the breaking of only a small number of hydrogen bonds. Further investigations aimed at resolving issues concerning the relationships of exchange chemistry to

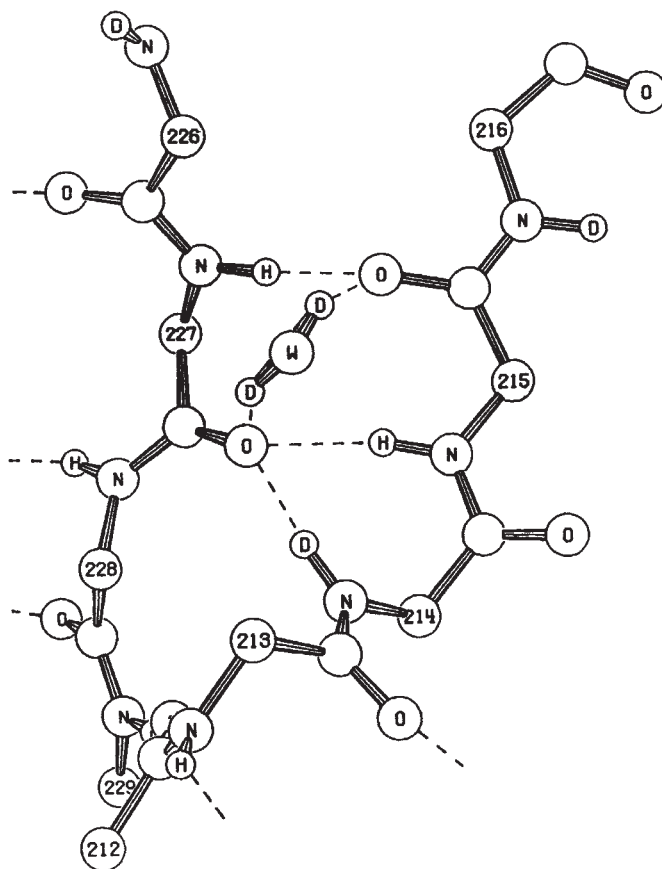


Fig. 7 Section of β -sheet showing the H-bond interaction between 227 NH-215 O, and 227 O-215 NH. An ordered D_2O molecule shares this H-bond, yet the peptide NH protons are unexchanged. The locations of the D atoms on the D_2O molecule were determined from the neutron density map, to indicate how the D_2O was oriented relative to the coordination ligands (215 O and 227 O). Note that the 214 amide is fully exchanged. Calculations indicate that the 214 and 215 amide protons are equally close to the solvent interface. Their respective hydrogen bond strengths are quite different. The 227 O-215 N hydrogen bond is classified as normal (2.9 Å), whereas the 227 O-214 N distance is quite long (3.5 Å), indicating a weak interaction.

protein dynamics should profit by the application of modelling studies using the exchange information reported here as a data base.

I thank Dr S. Spencer for his assistance with all aspects of the trypsin analysis, J. Shprungin for technical help and Drs A.

Kossiakoff and G. Rose for advice during the preparation of the manuscript. The stereo space filling drawings (Fig. 3) were produced at the NIH by Richard Feldman. Research was carried out at Brookhaven National Laboratory under the auspices of the United States Department of Energy.

Received 28 October 1981; accepted 15 February 1982.

1. Linderstrom-Lang, K. U. & Schellman, J. A. in *The Enzymes* Vol. 1, 2nd edn (eds Boyer, D., Lardy, H. & Myrback, K.) 443 (Academic, New York, 1959).
2. Gurd, F. N. & Rothgeb, T. M. *Adv. Protein Chem.* **33**, 73 (1979).
3. Sternberg, M. J., Grace, D. E. & Phillips, D. C. *J. molec. Biol.* **130**, 231 (1979).
4. Frauenfelder, H., Petsko, G. & Tsernoglou, D. *Nature* **280**, 558 (1979).
5. McCammon, J. A., Gelin, B. R. & Karplus, M. *Nature* **267**, 585 (1977).
6. Lumry, R. & Biltonen, R. in *Structure and Stability of Biological Macromolecules*, 65 (Dekker, New York, 1969).
7. Baldwin, R. L. A. *Rev. Biochem.* **44**, 453 (1975).
8. Blumenfeld, L. A. *J. theor. Biol.* **58**, 269 (1976).
9. Hvidt, A. & Linderstrom-Lang, K. *Biochim. biophys. Acta* **14**, 574 (1954).
10. Hvidt, A., Johansen, G. & Linderstrom-Lang, K. in *A Laboratory Manual of Analytical Techniques in Protein Chemistry* (eds Alexander, P. & Block, R. J.) 101 (Pergamon, New York, 1960).
11. Hvidt, A. & Nielsen, S. O. *Adv. Protein Chem.* **21**, 287 (1966).
12. Englander, S. W., Downer, N. W. & Teitelbaum, H. A. *Rev. Biochem.* **41**, 903 (1972).
13. Englander, S. W. *et al. Biophys. J.* **32**, 557 (1980).
14. Woodward, C. K. & Hilton, B. D. *Rev. Biophys. Bioengng* **8**, 99 (1979).
15. Lumry, R. & Rosenberg, A. *Colloques int. Cent. natn. Rech. scient.* **246**, 53 (1975).
16. Wüthrich, K. & Wagner, G. *J. molec. Biol.* **130**, 1 (1979).
17. Wagner, G. & Wüthrich, K. *J. molec. Biol.* **134**, 75 (1979).
18. Hilton, B. D. & Woodward, C. K. *Biochemistry* **17**, 3325 (1978).
19. Schreier, A. A. & Baldwin, R. L. *J. molec. Biol.* **105**, 409 (1976).
20. Rosa, J. J. & Richards, F. M. *J. molec. Biol.* **133**, 399 (1979).
21. Crick, F. H. C. & Kendrew, J. C. *Adv. Protein Chem.* **12**, 133 (1957).

22. Richards, F. M. A. *Rev. Biochem.* **32**, 269 (1963).
23. Rupley, J. A. in *Structure and Stability of Biological Macromolecules*, 291 (Dekker, New York, 1969).
24. Matthews, B. W. *J. molec. Biol.* **33**, 491 (1968).
25. Perutz, M. F. *Trans. Faraday, Soc.* **B42**, 187 (1946).
26. Low, B., Richards, F. M. & Berger, J. E. *J. Am. chem. Soc.* **78**, 1107 (1956).
27. Haggis, G. H. *Biochim. biophys. Acta* **23**, 494 (1957).
28. Praissman, M. & Rupley, J. A. *Biochemistry* **7**, 2431 (1968).
29. Praissman, M. & Rupley, J. A. *Biochemistry* **7**, 2446 (1968).
30. Tuchsén, E., Hvidt, A. & Ottesen, M. *Biochimie* **62**, 563 (1980).
31. Kossiakoff, A. A. & Spencer, S. A. *Biochemistry* **20**, 6462 (1981).
32. Alberi, J. L. *Brookhaven Symp. Biol.* **27**, VIII-24 (1975).
33. Spencer, S. A. & Kossiakoff, A. A. *J. appl. Crystallogr.* **13**, 563 (1980).
34. Chambers, J. L. & Stroud, R. M. *Protein Data Bank* (Brookhaven National Laboratory, 1977).
35. Salemme, F. R. *J. molec. Biol.* **146**, 143 (1981).
36. Molday, R. S., Englander, S. W. & Kallen, R. G. *Biochemistry* **11**, 150 (1972).
37. Kim, P. S. & Baldwin, R. L. *Biochemistry* **21**, 1 (1982).
38. Klotz, I. M. & Frank, B. H. *J. Am. chem. Soc.* **87**, 2721 (1965).
39. Scarpa, J. S., Mueller, D. D. & Klotz, I. M. *J. Am. chem. Soc.* **89**, 6024 (1967).
40. Brewster, A. I. & Bovey, F. A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 1199 (1971).
41. Hvidt, A. & Covett, R. *J. Am. chem. Soc.* **92**, 5546 (1970).
42. Chou, P. Y. & Fasman, G. D. A. *Rev. Biochem.* **47**, 251 (1978).
43. Artymuk, P. J. *et al. Nature* **28**, 563 (1979).
44. Knox, D. & Rosenberg, A. *Biopolymers* **19**, 1049 (1980).
45. Hilton, B. D., Trudeau, K. & Woodward, C. K. *Biochemistry* **20**, 4697 (1981).
46. Birktoft, J. J. & Blow, D. M. *J. molec. Biol.* **68**, 187 (1972).

LETTERS TO NATURE

The formation of galactic haloes in the neutrino-adiabatic theory

Adrian L. Melott*

Center for Relativity Theory, The University of Texas at Austin, Texas 78712, USA

Recent evidence of a finite rest mass for the neutrino has led to a revival of the idea that these particles may provide the unseen but dynamically indicated matter on cosmological scales of galactic halo and larger size. In the early Universe, the free streaming of collisionless neutrinos is seen to damp out perturbations less than of the order of tens of Mpc long¹⁻³. A numerical simulation of the clustering of collisionless particles in a universe with this type of perturbation spectrum⁶ has led to the formulation of a cell structure with large voids. Many observations^{7,8} support the existence of such a structure on scales of tens of Mpc. However, collapse of perturbations of this size leads to velocities of the order of 1,000 km s⁻¹, and as the neutrinos presumably cannot cool, this seems to be too large for their inclusion in galactic haloes, which have much smaller velocity dispersions. Thus it has been suggested that neutrinos cannot cluster in galactic haloes without some means to 'save' smaller-scale perturbations^{3,4,9}. Furthermore, if neutrinos of mass ~30 eV form galactic haloes, the phase space density must be very close to the primordial one¹⁰—and phase mixing is often associated with collapse. Although the average value of the velocity dispersion may be reduced if the perturbations are anisotropic¹¹, a disparity still remains. I now present a numerical simulation which shows that if the initial perturbations are sufficiently anisotropic, the collapse of very large perturbations of collisionless particles leads to a condensation of particles with low velocity dispersion and high phase-space density, which may easily fragment.

A series of numerical simulations were initiated to study the collapse of large-scale perturbations of massive neutrinos. The technique used was the "cloud in cell" method¹², which accurately models collective effects while suppressing two-body effects (such as the large-angle gravitational scattering of one neutrino by another), which are not important for neutrino clustering. The gravitational potential was derived using a Fast Fourier Transform method, with periodic boundary conditions, in co-moving coordinates⁶. I began with a study of the plane-symmetric case, an approximation to highly anisotropic collapse, because anisotropy may increase during collapse¹¹.

A high-resolution one-dimensional simulation, to be published elsewhere¹⁴, showed that the central region of the phase space (see Fig. 1) is dominated by a high phase-space density, low velocity dispersion condensation. A sinusoidal perturbation of 100 Mpc wavelength was chosen for the simulation so as to test the hypothesis severely. The 100 Mpc field was divided in successive simulations into 10², 10³ and 10⁴ grids, onto which were placed 10³, 10⁴ and 10⁴ particles respectively. Changing the resolution from slightly larger to much smaller than the condensation had no effect on the conclusion, apart from low-level noise introduced in the third simulation by the smaller number of clouds per cell. Imposition of even highly asymmetric initial perturbations resulted in an asymmetric condensation, with a small net momentum, but still with the crucial high phase-space density and low velocity dispersion.

For a sinusoidal initial perturbation of amplitude $\delta\rho/\rho \sim 10^{-3}$ at $z = 10^4$ and a wavelength of 100 Mpc, in a universe of Hubble constant 75 km s⁻¹ Mpc⁻¹ and Ω (density relative to critical density) of 1.07, about 12% of the (10⁴) particles were bound in the central Mpc, with velocities < 200 km s⁻¹. Furthermore, the condensation was found to have a phase-space density very close to that of the primordial distribution.

A two-dimensional code was developed, using a 100 × 100 grid field scaled to 1 Mpc per grid. This code was found to conserve energy to 1.7% or better (a considerable improvement over past calculations of this type⁶), depending on choice of time step; to reverse itself from a collapsed state to the uniform distribution; and to behave like an analytical solution for the

* Present address: Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.

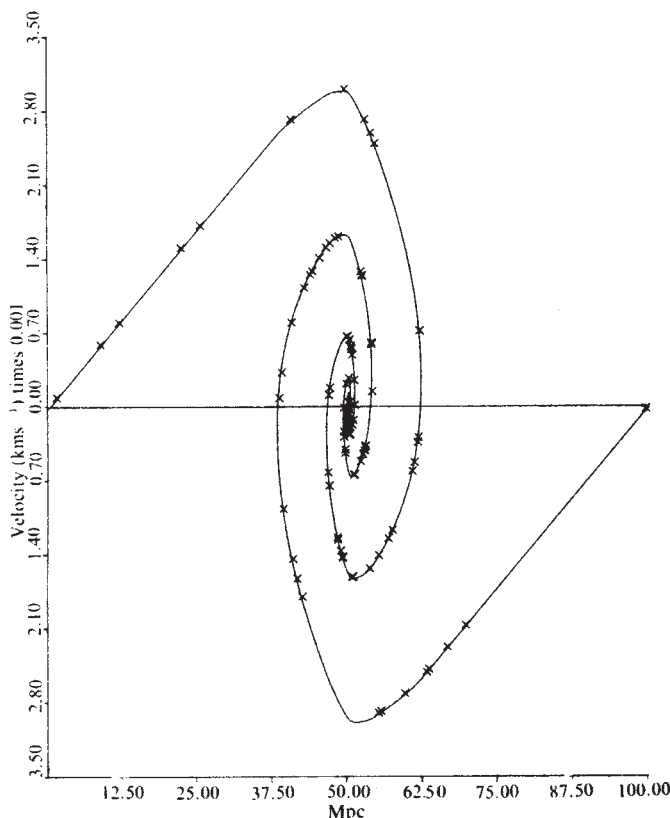


Fig. 1 A plot of the phase space of the one-dimensional simulation at $Z=0$. A solid line connects the particles, and a cross is drawn at the position of every one-hundredth test particle. Note the condensation of particles to the central Mpc.

a condensing protogalaxy can capture a significant number of neutrinos, resulting in a density enhancement in a central region, so that the neutrino density becomes comparable with the baryon density, which is demanded by flat galactic rotation curves.

Therefore the collapse of large-scale perturbations of collisionless particles results in a condensation of particles of low velocity dispersion and high phase-space density. This con-

case of uniform spherical symmetric perturbations. It contained two populations of particles: those given no initial velocity and a test population with thermal velocities appropriate to 30 eV neutrinos.

The result of collapse to $Z=0$ is seen in Fig. 2. A central slab is visible in the plot in which the darkness of a grid is proportional to its mass density. An $x-V_x$ projection of this result duplicates Fig. 1. The central region would be the predicted site of galaxy formation in the adiabatic theory, as the 'normal' matter undergoes collision and shocking. Somehow the collisionless neutrinos must cluster around these galaxies.

Figure 3 shows the result of assigning only 1% of the mass to 'protogalaxy grids' at $Z=5$, the moment of pancaking of the neutrinos. The pattern again is that of the distribution of test particles, which are strongly clumped around the protogalaxies. The same clumping appeared if the mass was slowly assigned to protogalaxies from $Z=5$ to 4. In this case, it was found that 7% of the test particles were bound within the protogalaxy grids and had velocities $< 200 \text{ km s}^{-1}$. If the protogalaxies are chosen to be of smaller mass, distributed more evenly, the amount of material 'captured' increases because fewer particles acquire velocities greater than the 200 km s^{-1} limit. Very small perturbations are sufficient to fragment the neutrino pancake.

Figures 2 and 3 are not to the same gray scale—black is adjusted to be the grid of highest density. The dark grids in Fig. 3 are of the same density as the local group, if it is bound.

Therefore sufficiently anisotropic perturbations will permit massive neutrinos to cluster in regions of low velocity dispersion, high mass density and high phase-space density. P. J. E. Peebles (personal communication) has pointed out that the central mass density of these regions must be much larger than the average shown here. My simulation cannot accurately model the dynamics inside a grid. It has been shown¹³, however, that

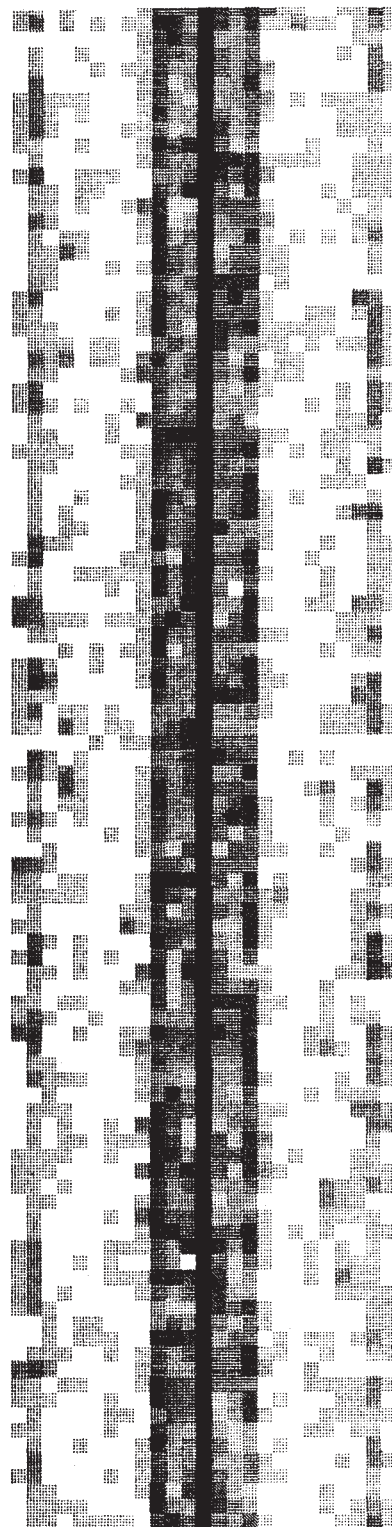


Fig. 2 A plot of the density of the two-dimensional simulation, at $Z=0$, given the same initial conditions as in the one-dimensional simulation. A small population of particles exists outside the slab, too light to be seen in this rendering.

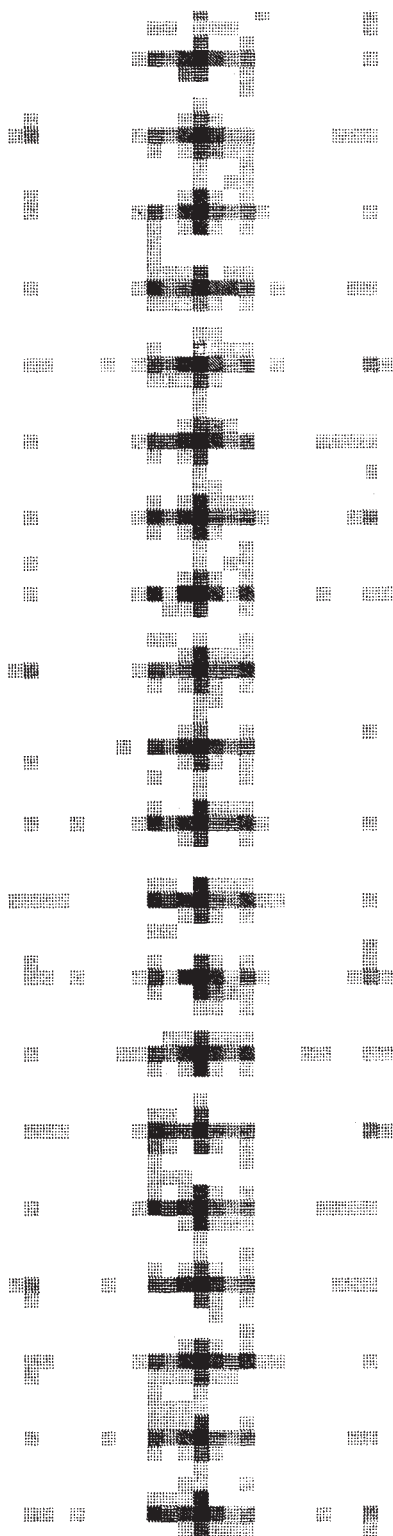


Fig. 3 A plot of the same initial conditions as in Fig. 2, if 1% of the mass is divided between 'protogalaxies' in the caustic surface. The clustering of the test particles is pronounced.

densation is subject to fragmentation by introduction of small 'protogalaxy' mass perturbations in the central plane, and is available for capture by collapsing protogalaxies.

Since many neutrinos would lie beyond the region of galaxy clustering and the Hubble expansion rate may be highly anisotropic, estimates of Ω from observations of the deceleration of galaxies may be incorrect. Therefore the clustering of massive neutrinos to account for the 'hidden mass' on all scales

is not in conflict with simple assumptions, provided that the initial perturbations are sufficiently anisotropic.

I thank D. W. Sciama for support, encouragement and training, S. Weinberg for financial support, and the Department of Physics, University of Texas at Austin for allowing the large amount of computer time needed to conduct this study. A conversation with J. Binney was helpful in planning the study, and N. Sharp assisted in computing.

Received 23 November 1981; accepted 17 February 1982.

1. Bisnovatyi-Kogan, G. S. & Novikov, I. D. *Soviet Astr.* **24**, 516-517 (1980).
2. Doroshkevich, A. G., Zel'dovich, Ya. B., Syunyaev, R. A. & Klopov, M. Yu. *Soviet astr. Lett.* **6**, 252-259 (1980).
3. Bond, J. R., Efstathiou, G. & Silk, J. *Phys. Rev. Lett.* **45**, 1980-1984 (1980).
4. Klinkhauser, F. R. & Norman, C. *Astrophys. J. Lett.* **243**, L1-L4 (1981).
5. Wasserman, I. *Astrophys. J.* **248**, 1-12 (1981).
6. Doroshkevich, A. G. et al. *Mon. Not. R. astr. Soc.* **192**, 321-337 (1980).
7. Einasto, J., Joëveer, M. & Saar, E. *Mon. Not. R. astr. Soc.* **193**, 353-375 (1980).
8. Kirshner, R. P., Oemler, A. Jr., Schechter, P. L. & Schechtman, S. A. *Astrophys. J. Lett.* **248**, L57-L60 (1981).
9. Davis, M., Lecar, M., Pryor, C. & Witten, E. *Astrophys. J.* **250**, 423-431 (1981).
10. Tremaine, S. & Gunn, J. L. *Phys. Rev. Lett.* **42**, 407-410 (1979).
11. Barrow, J. & Silk, J. *Astrophys. J.* **250**, 432-439 (1981).
12. Hockney, R. W. & Eastwood, H. W. *Computer Simulation Using Particles* (McGraw Hill, New York, 1981).
13. Zel'dovich, Ya. B., Klypin, A. A., Khlopov, M. Yu. & Chechetkin, V. M. *Sov. J. nucl. Phys.* **31**, 664-669 (1980).
14. Melott, A. *Phys. Rev. Lett.* (in the press).

High-energy γ -ray light curve of PSR0531+21

R. D. Wills*, K. Bennett*, G. F. Bignami†, R. Bucccheri‡, P. A. Caraveo†, W. Hermsen§, G. Kanbach||, J. L. Masnou¶, H. A. Mayer-Hasselwander||, J. A. Paul¶ & B. Sacco‡

* Space Science Department of the European Space Agency, ESTEC, Noordwijk, The Netherlands

† Istituto di Fisica Cosmica del CNR, Milano, Italy

‡ Istituto di Fisica Cosmica e Informatica del CNR, Palermo, Italy

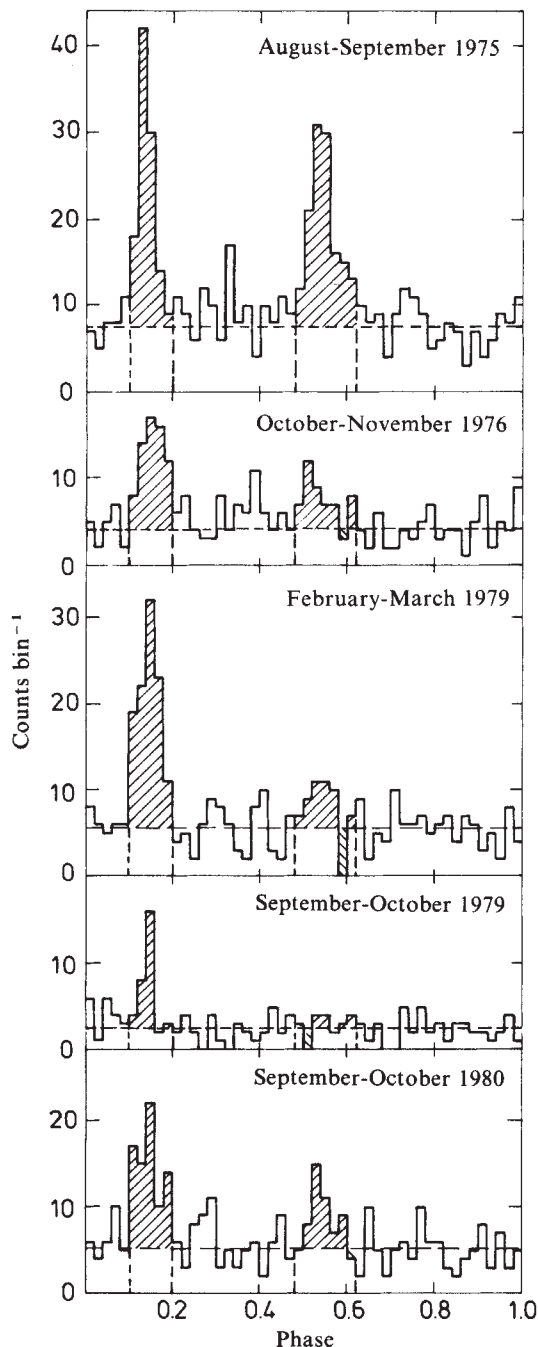
§ Cosmic Ray Working Group, Huygens Laboratorium, Leiden, The Netherlands

|| Max Planck Institut für Extraterrestrische Physik, Garching bei München, FRG

¶ Service d'Electronique Physique, Centre d'Etudes Nucléaires de Saclay, France

The Crab pulsar (PSR0531+21) was one of the earliest identified sources of high-energy γ radiation¹⁻⁶. As such it was selected as the first object of detailed study by ESA's γ -ray astronomy satellite COS B, launched in August 1975. The experiment and mission have been described elsewhere⁷. In the first 6 yr of operation COS B has made five observations (each of 30-40 days duration) in which PSR0531+21 was within 15° of the centre of the field of view. The first of these confirmed the result of the SAS 2 experiment that the γ -ray light curve for energies above a few tens of MeV consisted essentially of two equally strong peaks but the later measurements found a decrease in the strength of the second peak relative to the first. No change was observed in the 2-12 keV X-ray light curve simultaneously measured by COS B. The total data from the five observations contain the first evidence of interpulse emission between the two peaks in the energy range 50-3,000 MeV.

The epochs and observing conditions are summarized in Table 1. The accuracy of the on-board clock and the precision of reconstitution of the position of the satellite in its orbit as a function of time permit the measured arrival times of the γ -ray photons to be transformed to the Solar System barycentre (SSB) with an uncertainty of 0.5 ms. In making this transformation the pulsar position measured at radio frequencies⁸ was used.



The phase ϕ of a γ -ray event with SSB arrival time T is calculated from the frequency f and its derivatives $\dot{f}$, $\ddot{f}$ at epoch T_0 according to

$$\phi = \phi_0 + \Delta T f + \frac{1}{2} \Delta T^2 \dot{f} + \frac{1}{6} \Delta T^3 \ddot{f}$$

where

$$\Delta T = T - T_0$$

The values of these parameters used to analyse the present observations are given in Table 1. For the first three observations radio parameters were available (ref. 9 and personal communications from V. Boriakoff and D. Ferguson, and J. M. Rankin). For observation 54 use was made of the COS B 'pulsar synchronizer,' a proportional counter sensitive to 2–12 keV X rays¹², the arrival times of which are measured with the same precision as for the γ rays. As there are many more X-ray photons than γ rays, it was possible to determine the pulsar frequency and its first derivative by testing trial light curves for values of these parameters scanned around the extrapolations from earlier epochs. The values adopted are those which gave

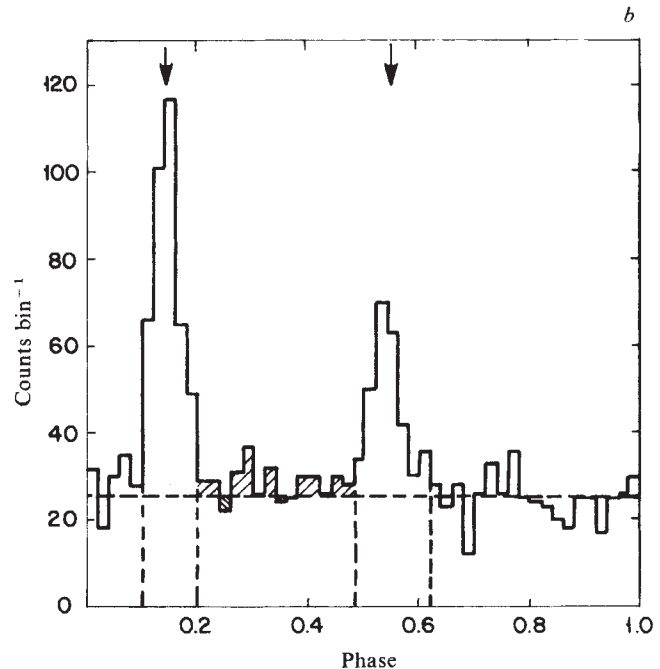


Fig. 1 *a*, γ -ray light curves of PSR0531+21 at the epochs indicated. Events were selected by energy E and arrival direction θ according to $50 < E < 150$ MeV and $\theta < 7^\circ$ or $150 < E < 3,000$ MeV and $\theta < 4^\circ$. The shaded areas indicate the phase intervals of the two pulses and the horizontal broken lines show the level of the background. *b*, γ -ray light curve obtained by summing the individual curves. The horizontal broken line shows the background level and the shaded area indicates the interpulse emission. The two arrows indicate the phases of the radio peaks.

the X-ray light curve having the highest χ^2 . (For a discussion of the method see ref. 11.) This approach could not be used for observation 44 because the pulsar was outside the field of view of the X-ray detector. Instead the arrival times of the γ rays themselves were used in a parameter scan. The features of the resulting light curve may therefore be over-emphasized but this observation is trivial in the final conclusions.

High-quality γ -ray events were selected from the data by applying established and calibrated criteria to the spark-chamber pictures and counter pulse-height measurements. The energy spectra of the pulsar γ rays¹² and of the background radiation are such that the best signal-to-noise ratio is obtained within the energy range 50–3,000 MeV. To take account of the variation with energy of the instrument response (effective sensitive area and angular resolution) the data were analysed in two separate energy ranges, 50–150 MeV and 150–3,000 MeV. Within each range the signal-to-noise ratio could be further optimized by selecting events for which the measured arrival direction fell within a limiting cone centred on the direction of the pulsar. The half angle of the optimum cone was found to be 7° in the lower energy range and 4° in the upper. Figure 1*a* shows the γ -ray light curve for each observation obtained by summing the respective curves for the two ranges. These curves show the actual numbers of photons recorded in each phase bin. Absolute fluxes are not presented because their accurate determination depends on the analysis of the temporal development of the instrument's sensitivity over the five years and this analysis is still in progress.

Figure 1*a* shows clearly that for the later observations the second pulse in the light curve is weaker relative to the first than it is in the first observation. For a more quantitative study phase boundaries were assigned to the two peaks, namely 0.10–0.20 and 0.48–0.62 respectively. A background level, above which 'pulsed counts' were measured, was derived as the mean level in the phase interval 0.62–0.10. This background includes not only the instrumental background, the general

Table 1 COS B observations of PSR0531 + 21

Observation no.	1	14	39	44	54
Date	17 August– 17 September 1975	30 September– 2 November 1976	22 February– 3 April 1979	29 August– 10 October 1979	4 September– 17 October 1980
Pointing direction (<i>l, b</i>)	184°, –6°	195°, +4°	190°, 0°	172°, –12°	188°, –3°
Aspect angle of pulsar	1°	15°	8°	14°	5°
Parameters of analysis					
Epoch	2442647.0	2442982.0625	2443946.442	2444128.5	2444508.5
f (s ^{–1})	30.13744554816328	30.12634886	30.094467179	30.088458052	30.075925066
f (10 ^{–10} s ^{–2})	–3.8348446531	–3.8310414	–3.82125043	–3.81956	–3.81527838
f (10 ^{–20} s ^{–3})	1.1413101714	0.3956	1.225	1.224	1.22
Offset ϕ_0	0.352344	0.376843	0.958065	0.86	0.85
Pulse–strength ratio P_2/P_1					
50–150 MeV	1.34 ± 0.38	0.33 ± 0.27	0.36 ± 0.18	*	0.43 ± 0.32
150–3,000 MeV	0.74 ± 0.21	0.57 ± 0.31	0.35 ± 0.18	0.38 ± 0.27	0.31 ± 0.18
50–3,000 MeV	1.04 ± 0.20	0.49 ± 0.20	0.36 ± 0.12	0.48 ± 0.31	0.37 ± 0.17

* Not determined due to insufficient statistics.

galactic emission and any isotropic radiation but also any continuous (unpulsed) emission from the pulsar itself or the Crab nebula. The two peaks and the background levels are indicated in Fig. 1a.

Added together, the light curves of Fig. 1a yield Fig. 1b which represents the best available database on the high-energy γ -ray emission from the Crab pulsar and contains the first positive evidence in the energy range 50–3,000 MeV for high-energy γ -ray emission in the 'interpulse region' between the two peaks of the PSR0531 + 21 light curve. Averaged over all the observations this amounts to $(15 \pm 4)\%$ of the total pulsed emission (that is the phase interval 0.10–0.62) in the energy range 50–3,000 MeV. This is not inconsistent with the upper limit of 15% derived earlier from the first observation alone¹³ and is about half of the corresponding contribution in the case of PSR0833 – 45 (ref. 14). Within the wide limits of error, there is no evidence for any variation with time in the relative flux in the interpulse region. An evaluation of the phase difference between the two peaks in Fig. 1b gives 13.0 ± 0.5 ms, in agreement with corresponding measurements at radio, optical and X-ray wavelengths^{15,16}. It is also possible to measure the peak widths which are found to be 1.6 ± 0.4 ms (FWHM) and 2.0 ± 0.5 ms (FWHM) respectively.

The light curves shown in Fig. 1 represent only a fraction of the photons from the pulsar, that is those with measured arrival directions, θ , (relative to the pulsar direction) within the selection cones. To obtain the best estimate of the total numbers of photons the distributions of their arrival directions were fitted with the point-spread function describing the instrument's angular resolution appropriate to the energy range and to the aspect angle, α , of the observation. This point-spread function is approximated by the expression¹⁷

$$h(\theta) = 2\pi\theta \exp(-(\theta/\theta_0)^2c)$$

where $c = 0.8$ for $50 < E < 150$ MeV and 0.5 for $150 < E < 3,000$ MeV. The value of θ_0 decreases with aspect angle in the lower energy range, from 4.8° at $\alpha = 0^\circ$ to 4.3° at $\alpha = 15^\circ$ and increases in the higher range, from 1.2° at $\alpha = 0^\circ$ to 1.75° at $\alpha = 15^\circ$.

Separate fits were made for each observation in each energy range and for the photons having phases within each of the two peaks. By integrating the fitted curves the total numbers of events contributing to each peak, P_1 and P_2 respectively, were determined. The ratio of the strengths of the two peaks was calculated and the results are presented in Table 1. These show that during the first observation the second peak was significantly stronger (relative to the first peak) than at later times. The effect is present in both energy ranges but is more pronounced below 150 MeV. Using a χ^2 analysis it has been found that the chance probability of the total effect is $\sim 5\%$.

The results presented here show no evidence for any change in the γ -ray light curve between 1976 and 1980. The departure from the mean lies in the earlier (1975) measurement. That result is consistent with the light curve measured by SAS 2 at energies above 35 MeV in 1972 and 1973¹⁸. It is estimated that, taking phase intervals as close as possible to those used in analysing the COS B data, the ratio of the second pulse to the first was 1.33 ± 0.39 at the time of the SAS 2 observation. This ratio is plotted together with those from COS B in Fig. 2. The probability that all these results represent chance fluctuations about a common mean is $< 1\%$ combining a run test with the χ^2 analysis.

In a preliminary presentation of these results¹⁹ the γ -ray light curves for the Vela pulsar PSR0833 – 45 were also given for a similar time interval. No variation in the shape of the light curves was observed for that pulsar.

The measured variation in the Crab pulsar light curve cannot be ascribed to a change in the sensitivity of the instrument unless that change was energy dependent and the two pulses have markedly different energy spectra. Bennett *et al.*¹³ found no evidence for spectral differences sufficient to account for the effect seen here. Furthermore, a possible sensitivity change is not sufficiently energy dependent to influence the measured light curve of PSR0833 – 45¹⁹, for which the second pulse is known to have a harder spectrum than the first¹⁴.

The pulsar synchronizer on board COS B has allowed the long-term behaviour of the 2–12 keV X-ray light curve of the pulsar to be studied. Figure 3 shows the results for the three observations when the pulsar was within the field of view of the proportional counter. The shapes of these light curves exhibit excellent reproducibility. The same situation applies at

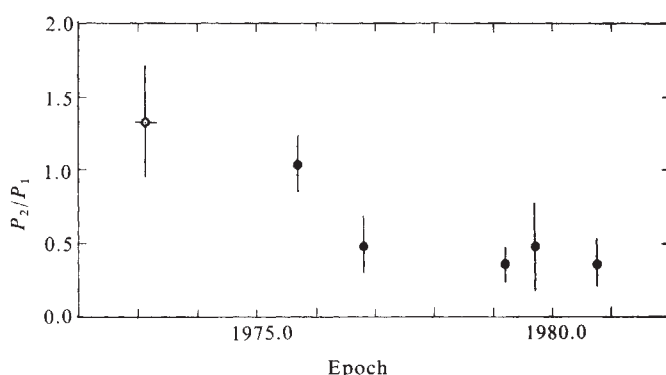


Fig. 2 Variation with time of the fraction of pulsed counts in the second pulse: $\circ$, > 35 MeV from SAS 2 data¹⁸; $\bullet$, 50–3,000 MeV from COS B data (present work).

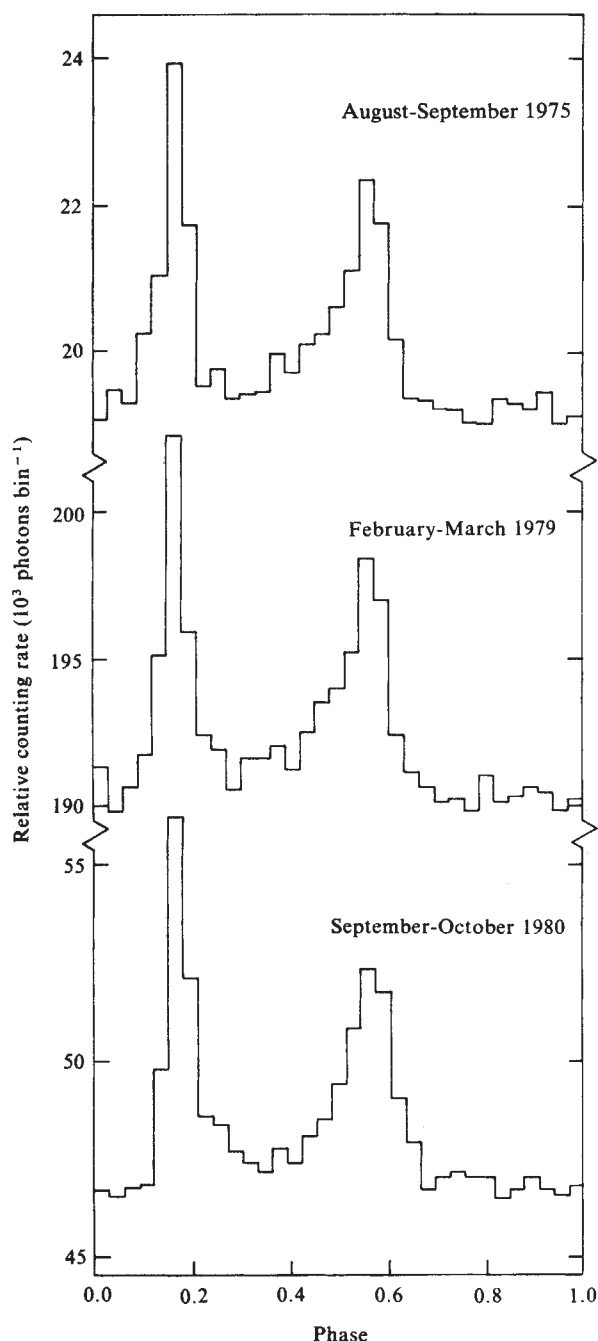


Fig. 3 X-ray (2–12 keV) light curves measured during observations 1, 39 and 54.

optical wavelengths where very small limits have been placed on the variation of the light curve²⁰. Above ~20 keV a decrease in the relative strength of the second X-ray peak between observations in 1969–70^{21–23} and in 1974²⁴ has been claimed²⁵. However, the effect is probably due to differences in the definitions of the phase intervals of the two peaks by the original investigators. In low energy (1–20 MeV) γ rays light curves of the same shape were measured in October 1977 and May 1979²⁶ but over this limited time span there is no evidence in the COS B result for any variation. In the ultra-high energy range (>100 GeV) a variation in the strengths of different components of the light curve between 1971 and 1973 has been reported²⁷.

The observed variation in the high-energy light curve could be due to a change in the emission modes in one of the polar regions of the magnetosphere of the neutron star, but also to a change in the geometry, such as a decrease in the angle

between the magnetic and rotational axes. Whatever the explanation, the absence of the effect in the X-ray light curve points to different emission regions and/or production processes in the two energy ranges. The properties of the different components of the radio light curve support the presence of two beaming mechanisms^{28,29}. Different emission regions or beaming geometries have also been proposed to account for differences between the radio, optical and γ -ray light curves of PSR0833–45 (refs 28–30).

We thank Drs J. M. Rankin, V. Baroakoff and D. Ferguson for communicating radio data and Dr I. I. Shapiro for making available the MIT barycentric ephemeris.

Received 7 December 1981; accepted 24 February 1982.

1. Browning, R., Ramsden, D. & Wright, P. J. *Nature phys. Sci.* **232**, 99–101 (1971).
2. Alhath, P. *et al. Nature* **240**, 221–224 (1972).
3. Kinzer, R. L., Share, G. H. & Seeman, N. *Astrophys. J.* **180**, 547–549 (1973).
4. Partier, B. *et al. Nature phys. Sci.* **242**, 117–120 (1973).
5. McBreen, B., Ball, S. E. Jr, Campbell, M., Greisen, K. & Koch, D. *Astrophys. J.* **184**, 571–580 (1973).
6. Kniffen, D. A. *et al. Nature* **251**, 397–399 (1974).
7. Scarsi, L. *et al. Proc. 12th ESLAB Symp. ESA SP-124*, 3–11 (1977).
8. McNamara, B. J. *Publ. astr. Soc. Pac.* **83**, 491–492 (1971).
9. Gullahorn, G. T., Isaacman, R., Rankin, J. M. & Payne, R. R. *Astr. J.* **82**, 309–312 (1977).
10. Boella, G., Buccheri, R., Burger, J. J., Coffaro, P. & Paul, J. *Proc. 9th ESLAB Symp. ESA SP-106*, 353–362 (1974).
11. Buccheri, R., D'Amico, N., Scarsi, L., Kanbach, G. & Masnou, J. L. *Proc. 12th ESLAB Symp. ESA SP-124*, 309–316 (1977).
12. Lichti, G. G. *et al. Non-Solar Gamma Rays, Advances in Space Exploration Vol. 7* (eds Cowsik, R. & Wills, R. D.) 49–53 (Pergamon, Oxford, 1980).
13. Bennett, K. *et al. Astr. Astrophys.* **61**, 279–284 (1977).
14. Kanbach, G. *et al. Astr. Astrophys.* **90**, 163–169 (1980).
15. Papaliolios, C., Carleton, N. P. & Horowitz, P. *Nature* **228**, 445–450 (1970).
16. Kestenbaum, H. L., Ku, W., Novick, R. & Wolff, R. S. *Astrophys. J. Lett.* **203**, L57–L61 (1976).
17. Hermesen, W. thesis, Univ. Leiden (1980).
18. Thompson, D. J., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Lamb, R. C. *Astrophys. J.* **213**, 252–262 (1979).
19. Wills, R. D. *et al. Proc. 17th int. Cosmic Ray Conf., Paris* **1**, 22–25 (1980).
20. Jones, D. H. P., Smith, F. G. & Nelson, J. E. *Nature* **283**, 50–51 (1980).
21. Floyd, F. W., Glass, I. S. & Schnopper, H. W. *Nature* **224**, 50–51 (1969).
22. Kurfess, J. D. *Astrophys. J. Lett.* **168**, L39–L42 (1971).
23. Smathers, H. W., Chubb, T. A. & Sadeh, D. *Nature phys. Sci.* **232**, 120–121 (1971).
24. Ryckman, S. G. *et al. Nature* **266**, 431–433 (1977).
25. Toor, A. & Seward, F. D. *Astrophys. J.* **216**, 560–564 (1977).
26. Graser, U. & Schönfelder, V. *Astrophys. J.* (submitted).
27. Grindlay, J. E., Helmken, H. F. & Weekes, T. C. *Astrophys. J.* **209**, 592–601 (1978).
28. Manchester, R. N. & Lyne, A. G. *Mon. Not. R. astr. Soc.* **181**, 761–767 (1977).
29. Arons, J. *Proc. IAU Symp. No. 94*, 175–204 (1981).
30. Buccheri, R. *Proc. IAU Symp. No. 95* 241–250 (1981).

Was the 1970 geomagnetic jerk of internal or external origin?

S. R. C. Malin & B. M. Hodder

Institute of Geological Sciences, West Mains Road, Edinburgh EH9 3LA, UK

Annual mean data from worldwide magnetic observatories show that there was a jerk¹ (step-change in the second time derivative) in the geomagnetic field, which took place over an interval of <2 yr around 1970. If such a short-lived phenomenon originated in the Earth's core, and were still detectable after passing through the mantle to the surface, this would imply a much lower conductivity for the lower mantle than had been believed previously. Here we show by an objective test that most of the jerk has an internal origin.

The geomagnetic field has a rich spectrum of temporal changes from audio frequencies to thousands of years. In general, the short-term variations (with time scales of a year or less) are of external origin relative to the Earth's surface, though they have a smaller internal component due to electromagnetic induction in the conducting layers of the Earth. The long-term variations, with time scales of several decades, are of internal origin. At intermediate time scales the picture is less clear; there are certainly external variations with a period

of ~11 yr, associated with the sunspot cycle, and there are variations that occur over a few years which seem to be of internal origin. These short-term variations are of particular interest because, if they are real and originate from the Earth's core, they place constraints on the conductivity of the mantle through which they have to pass before being detected at the surface².

In 1967, using accepted values for the mantle conductivity, Currie³ deduced that no core signal with a period of ≤ 4 yr could penetrate the mantle and that all observed variations on a shorter time scale must be of external origin. Later, Alldredge⁴ raised this limit to 15 yr and suggested a correspondingly higher value of mantle conductivity. Other authors prefer lower⁵⁻⁷ conductivities and time scales.

The most convincing evidence for the lower values comes from the change in geomagnetic secular acceleration that occurred over an interval of 1 or 2 yr around 1970. This event was first reported by Courtillot *et al.*⁸ (although Mizuno had submitted an article nearly a year earlier⁹ describing similar effects for 1965 and 1974 restricted to the region of Japan) and has since been widely studied⁷⁻¹⁵. Although the effect is worldwide^{14,15}, it is most clearly seen in declination data from European observatories. See, for example, Fig. 1a, which shows a plot of 3 yr running means of declination observed at Eskdalemuir, Scotland, and Fig. 1b, the first derivative of these data (secular variation) obtained by differencing successive means. The secular acceleration is the slope of curve *b*. It is reasonably constant before 1970 but changes suddenly at about that date to a new value, at which it remains for the rest of the interval shown.

The reality of this phenomenon, its worldwide distribution and the short interval over which it occurs are firmly established. However, although most authors^{11,15} believe the effect to be of internal origin, their reasons for this are based on morphological arguments rather than objective measurement, and there are others^{16,17} who believe that all such short-period variations are of external origin. We present here an objective separation into parts of internal and external origin and compare their sizes.

Gauss¹⁸ originally developed the method of spherical harmonic analysis to determine whether the main geomagnetic field was of internal or external origin, and the method may equally well be applied to time derivatives of the geomagnetic field. Here, we apply it to the change of secular acceleration that occurred around 1970. The data are annual mean values of the geomagnetic north, east and vertical components from the 83 observatories that operated throughout the interval 1961.0–1978.0 (ref. 19). For each observatory and element, means were formed for the centre of each even year by forming $(x_{n-1} + 2x_n + x_{n+1})/4$, where x_n denotes the annual mean value for year n , giving eight bi-annual means, denoted (1)–(8). The pre-1970 secular acceleration was defined as $[(1) - (2) - (3) + (4)]/2$ and the post-1970 secular acceleration as $[(5) - (6) -$

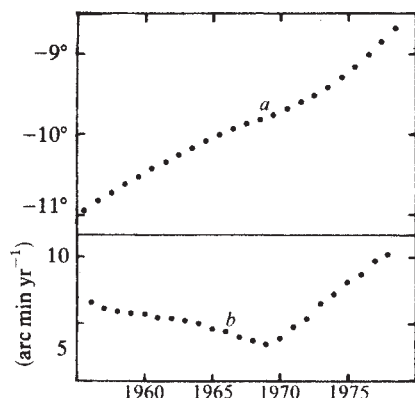


Fig. 1 Declination observed at Eskdalemuir, Scotland: *a*, mean values; *b*, secular variation obtained from first differences of the mean values. Note the sudden change of slope around 1970.

Table 1 Spherical harmonic coefficients of the change of secular acceleration that occurred around 1970

<i>m</i>	<i>n</i>	Internal		External	
		g_n^m (nT yr ⁻²)	h_n^m (nT yr ⁻²)	g_n^m (nT yr ⁻²)	h_n^m (nT yr ⁻²)
0	1	-2.44 ± 0.27	—	-1.48 ± 0.43	—
1	1	—	1.04 ± 0.29	—	—
0	2	1.03 ± 0.23	—	-0.78 ± 0.26	—
1	2	0.59 ± 0.22	-2.40 ± 0.26	—	-0.29 ± 0.26
2	2	-0.68 ± 0.27	—	—	-1.10 ± 0.29
0	3	1.04 ± 0.16	—	—	—
1	3	0.83 ± 0.18	-2.23 ± 0.20	—	0.94 ± 0.24
2	3	1.04 ± 0.20	-0.43 ± 0.20	—	—
3	3	—	—	—	-0.88 ± 0.24
0	4	—	—	—	—
1	4	0.28 ± 0.14	—	-0.28 ± 0.12	-0.20 ± 0.17
2	4	-0.77 ± 0.16	-0.41 ± 0.16	0.21 ± 0.15	0.35 ± 0.15
3	4	0.26 ± 0.16	0.95 ± 0.15	-0.35 ± 0.16	—
4	4	-0.94 ± 0.23	—	0.73 ± 0.24	-0.74 ± 0.24

All coefficients that exceed their standard deviation are included, up to maximum degree and order 4.

(7) + (8)]/2. The jerk¹ is simply the difference between these quantities.

In the analysis, it is assumed that the jerk may be adequately represented by the spherical harmonic expansion of a potential field up to maximum degree and order 4 for both internal and external terms. The spherical harmonic coefficients are obtained by the method of least squares, solving for internal and external coefficients simultaneously, and evaluating their standard deviations in the usual way from the sum of squares of residuals, number of degrees of freedom and the leading diagonal of the inverse matrix. After some experiment with different sets of coefficients, the solution including all those that exceed their standard deviations and omitting all those that do not was found to be that given in Table 1. Following recommended geomagnetic practice²⁰, the coefficients are in the Schmidt quasi-normalized form.

It is immediately obvious that most of the jerk is of internal origin: the mean square value²¹ of the internal part exceeds that of the external part by a factor of 3.4 ± 0.8 . Of the 10 largest coefficients, only 2 are for the external part. The larger of these, g_1^{0e} (using the notation of Chapman and Bartels²²), represents the field due to a ring of current above the Equator, which is precisely what would be expected for the external contribution. The only other external coefficient to exceed 1 nT yr^{-2} is h_2^{2e} ; there is no obvious external source which would produce this term. The three coefficients that exceed 2 nT yr^{-2} (g_1^{0i} , h_2^{2i} , h_3^{3i}) are all of internal origin. Part of the g_1^{0i} term could be due to induction by the current represented by g_1^{0e} , but, even if the whole Earth were a perfect conductor, this could not account for more than 0.74 nT of g_1^{0i} (ref. 23). We would not expect, and we do not find, any obvious pattern to the internal part, except the tendency¹⁵ for the coefficients to decrease in amplitude with increasing n . This tendency, which is more obvious when the coefficients are fully normalized, suggests that the jerk originates deep within the Earth rather than near the surface.

We conclude that internal sources can give rise to changes in secular variation on time scales as short as 1 or 2 yr and that these were indeed the major factor in the geomagnetic jerk that occurred around 1970.

This paper is published with the permission of the director of the Institute of Geological Sciences (NERC).

Received 19 January; accepted 5 March 1982.

- Schot, S. H. *Am. J. Phys.* **46**, 1090–1094 (1978).
- Runcorn, S. K. *Trans. Am. geophys. Un.* **36**, 191–198 (1955).
- Currie, R. G. *J. geophys. Res.* **72**, 2623–2633 (1967).
- Alldredge, L. R. *J. geophys. Res.* **82**, 5427–5431 (1977).
- Stacey, F. D., MacQueen, H. W., Smylie, D. E., Rochester, M. G. & Conley, D. *Trans. Am. geophys. Un.* **59**, 1027 (1978).
- Yukutake, T. *J. Geomagn. Geoelect.*, Kyoto **25**, 195–212 (1973).
- Achache, J., Courtillot, V., Ducruix, J. & Le Mouél, J.-L. *Phys. Earth planet. Inter.* **23**, 72–75 (1980).

8. Courtillot, V., Ducruix, J. & Le Mouél, J.-L. *C. r. heb. Séanc. Acad. Sci., Paris* **D287**, 1095–1098 (1978).
9. Mizuno, H. *J. Geomagn. Geoelect., Kyoto* **32**, 413–420 (1980).
10. Alldredge, L. R. *C. r. heb. Séanc. Acad. Sci., Paris* **B289**, 169–171 (1979).
11. Courtillot, V., Ducruix, J. & Le Mouél, J.-L. *C. r. heb. Séanc. Acad. Sci., Paris* **B289**, 173–175 (1979).
12. Ducruix, J., Courtillot, V. & Le Mouél, J.-L. *Geophys. J. R. astr. Soc.* **61**, 73–94 (1980).
13. Le Mouél, J.-L., Madden, T. R., Ducruix, J. & Courtillot, V. *Nature* **290**, 763–765 (1981).
14. Chan, H. D., Ducruix, J. & Le Mouél, J.-L. *C. r. heb. Séanc. Acad. Sci. Paris* **B293**, 157–160 (1981).
15. Malin, S. R. C., Hodder, B. M. & Barraclough, D. R. *Ebro Observatory 75th Anniversary Vol.* (in the press).
16. Alldredge, L. R. *J. geophys. Res.* **80**, 1571–1578 (1975).
17. Alldredge, L. R. *J. Geomagn. Geoelect., Kyoto* **32**, 421–425 (1980).
18. Gauss, C. F. *Allgemeine Theorie des Erdmagnetismus* (Leipzig, 1839).
19. Lewis, A. G. *Geomagn. Bull. Inst. geol. Sci.* **10**, 1–296 (1981).
20. Goldie, A. H. R. & Joyce, J. W. (eds) *Association of Terrestrial Magnetism and Electricity—Trans. Washington Meet.*, 550 (Neill, Edinburgh, 1940).
21. Lowes, F. J. *J. geophys. Res.* **71**, 2179 (1966).
22. Chapman, S. & Bartels, J. *Geomagnetism* 639 (Clarendon, Oxford, 1940).
23. Banks, R. *Geophys. J. R. astr. Soc.* **17**, 457–487 (1969).

Intensity of the geomagnetic field near Loyang, China between 500 BC and AD 1900

Q. Y. Wei, D. J. Li, G. Y. Cao, W. S. Zhang & S. P. Wang

Institute of Geophysics, Academia Sinica, Peking, China

There are remarkably few data on the intensity of the geomagnetic field in early China. New results reported here indicate that the total intensity of the magnetic field of the Earth in the region near Loyang (lat. 34°42' N, long. 112°40' E) has changed considerably during the past 2,400 yr (500 BC–AD 1900). The maximum value reached about 1,700 yr ago is as much as 54% higher than the present one.

Loyang is the ancient capital of nine dynasties, so samples collected from the city and nearby regions are suitable for archaeomagnetic research. The samples used in the present work are mainly of three types: (1) baked earth from ancient pottery or brick kilns and ovens in residential structures; (2) grave bricks—some of them are engraved with accurate dates and place of their baking; (3) bricks from ancient structures such as buildings and city wall.

Table 1 gives the relevant data on samples collected from Loyang and Zhengzhou. The stability of the primary remanent magnetization was tested in all the samples and proved satisfactory¹. The magnetization of the specimens was measured with an astatic magnetometer. The total intensity of the geomagnetic field was inferred from stepwise thermal demagnetization^{2,3} with temperature steps of 40–60 °C from 100 to ~600 °C. The slope of the natural remanent magnetization–partial thermal remanent magnetization (NRM–PTRM) straight line was calculated using the method of least squares. The points in the

NRM–PTRM diagram which deviated systematically at low and high temperatures from a straight line were discarded.

The averaged data for each dynasty were obtained by using the method we have described earlier⁴, and are given in Table 1. Figure 1 shows variation of the total intensity of the geomagnetic field during the past 2,400 yr. The horizontal error bars represent the archaeological estimate of the duration for each epoch (or dynasty); the vertical error bars represent the standard deviation.

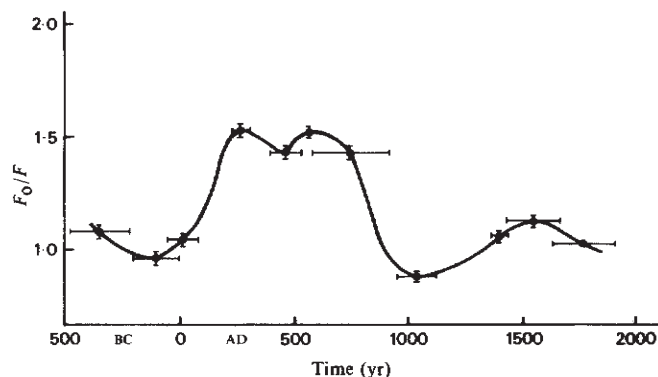


Fig. 1 The change in the ratio of the ancient intensity of the geomagnetic field, F_0 , to the present-day value F , in Loyang region.

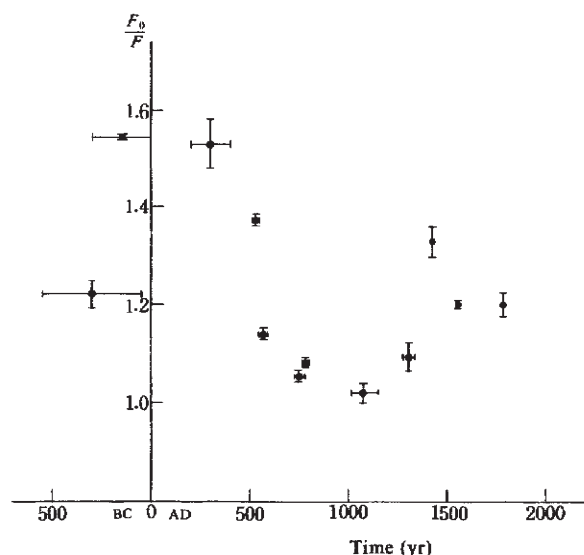


Fig. 2 Ratio of intensity of the ancient geomagnetic field (F_0) to that at present (F) in Japan³.

Table 1 Data of samples from Zhengzhou and Loyang

Sampling site	Lat. N	Long. E	Dynasty	Age	Sample type	N	F_0/F	F_0 (Oe)	δF
Zhengzhou	34°54'	113°39'	Zhanguo	475–221 BC	Grave brick	2	1.09	0.60	0.01
Loyang	34°42'	112°40'	Western Han	206 BC–AD 8	Baked earth	4	0.96	0.53	0.01
Loyang	34°42'	112°40'	Han (Wang Mon)	AD 9–23	Grave brick	5	1.05	0.58	0.01
Loyang	34°42'	112°40'	Wei-Jin	AD 220–300	Brick, baked earth	2	1.54	0.85	0.01
Loyang	34°42'	112°40'	Northern Wei	AD 386–534	Baked earth	3	1.43	0.79	0.02
Loyang	34°42'	112°40'	Northern Qi	AD 562–565	Grave brick	11	1.53	0.84	0.01
Loyang	34°42'	112°40'	Sui-Tang	AD 581–907	Brick, baked earth	10	1.43	0.79	0.02
Loyang	34°42'	112°40'	Northern Song	AD 960–1127	Brick, baked earth	8	0.88	0.48	0.01
Loyang	34°42'	112°40'	Early Ming	AD 1368–1435	Brick	2	1.06	0.58	0.01
Loyang	34°42'	112°40'	Ming	AD 1436–1661	Brick	2	1.13	0.62	0.01
Loyang	34°42'	112°40'	Qing	AD 1631–1911	Brick	4	1.03	0.57	0.01

F_0 , ancient intensity of the geomagnetic field; F , present-day intensity. N , number of samples.

It is clear from Fig. 1 that the total intensity of the geomagnetic field in the region near Loyang has changed. Between 400 BC and AD 1900 there are two minima, at about 100 BC and AD 1100, one (main) bimodal maximum between AD 250 and 600, and a smaller maximum at about AD 1550. The difference between the maximum ($\sim 85 \mu\text{T}$) and both minimum values ($\sim 53 \mu\text{T}$ and $\sim 48 \mu\text{T}$ respectively) amounts to about $37 \mu\text{T}$. The length of time between the two minima is about 1,200 yr.

It is interesting to compare these data with those obtained from other locations in the world. For example, comparison of the results from Loyang with those given in ref. 5 indicates that they are broadly similar although the details vary. Moreover the variation of the ratio of the total intensity in early Japan to that in present-day Japan³ is in a good agreement with that for Loyang both in tendency and magnitude of changes during the past 2,400 yr, although there is a time delay over the whole interval (see Fig. 2). This is also the case in comparisons with data from Athens⁶, Czechoslovakia and central America⁷. It may be related to the westward drift of the non-dipole field.

We thank Drs Xu Jing-yuan and Huang Ming-lang for supplying samples used in work on the secular variation of the total intensity as well as on the direction of the ancient geomagnetic field, and Professor J. A. Jacobs for reading the manuscript and for valuable comments.

Received 21 October 1981; accepted 19 February 1982.

1. Wei, Q. Y., Li, D. J., Cao, G. Y., Chang, W. S. & Wang, S. P. *Phys. Earth planet. Inter.* **25**, 107–112 (1981).
2. Thellier, E. & Thellier, O. *Ann. Geophys.* **15**, 285–376 (1959).
3. Nagata, T., Arai, Y. & Momose, K. *J. geophys. Res.* **68**, 5277–5281 (1963).
4. Wei, Q. Y., Li, D. J., Cao, G. Y., Chang, W. S. & Wang, S. P. *Acta geophys. sin.* **23**, 403–414 (1980).
5. Burlatskaya, S. P. *Izv. Akad. Nauk. S.S.S.R. Fys. Zem.* **4**, 5656 (1977).
6. Walton, D. *Nature* **277**, 643–644 (1979).
7. Bucha, V., Taylor, R. E., Berger, R. & Haury, E. W. *Science* **168**, 111 (1970).

A spinel to β -phase transformation mechanism in $(\text{Mg,Fe})_2\text{SiO}_4$

G. D. Price & A. Putnis

Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK

D. G. W. Smith

Department of Geology, University of Alberta, Edmonton, Alberta, Canada T6G 2E3

Olivine of approximate composition $\text{Mg}_{1.8}\text{Fe}_{0.2}\text{SiO}_4$ is a major constituent of the Earth's upper mantle. Structural changes, which occur in olivine as a function of increasing pressure and which involve the formation of the β -phase and spinel structure, γ -phase polymorphs, have therefore aroused considerable interest^{1–4}. The mechanisms of these transformations are of particular importance because of their effect on the properties of the mantle in the zone in which they occur. Several mechanisms have been proposed^{4–7}, but remain controversial in the absence of sufficient observational data on intergrowths of reactant and product phases. Although these phases have been produced experimentally², direct observations of such intergrowths have generally been made only on natural material from shock-produced veins in chondritic meteorites^{5,6,8}. Here we report observations made on the natural, high-density polymorphs of $(\text{Mg,Fe})_2\text{SiO}_4$ found in the Peace River meteorite, which reveal that the mechanism for the spinel to β -phase transformation involves the topotactic replacement of the spinel polymorph by single, β -phase grains.

We have previously described^{5,6} the microstructures in the spinel polymorph of $(\text{Mg,Fe})_2\text{SiO}_4$ (ringwoodite) from the Tenham chondritic meteorite, and suggested that the defects (Fig. 1 in ref. 5) found in the spinel were due to the partial inversion of ringwoodite to the β -phase. The defects found in the spinel lie on $\{110\}$ planes, and contrast analysis indicates^{6,8} that for a (110) defect, the displacement vector is of the type $a/4[110]$, or equivalently $a/4[1\bar{1}2]$ (refs 8,9). The defects are therefore stacking faults, which may be viewed either as Frank-type faults corresponding to the removal of a stoichiometric $(\text{Mg,Fe})_2\text{SiO}_4$ layer accompanied by an $a/4[110]$ displacement, or as Shockley-type faults corresponding to an $a/4[1\bar{1}2]$ shear. Stacking faults of this type leave the oxygen sub-lattice unaffected, but change the cation distribution so that in the fault plane the structure is that of the β -phase. The spinel and β -phase structures are closely related^{10,11}, and indeed the spinel structure can be completely transformed into that of the β -phase by the introduction of two $a/4[110]$ (110) stacking faults per unit cell. The transformation of spinel to β -phase may, therefore, be considered relatively simple, and one which would, in suitable conditions, be favoured kinetically over the reconstructive inversion to olivine. In this respect, our conclusions differ from those of Madon and Poirier⁸ who assume that the faults in ringwoodite result from the impingement of spinel domains nucleated at various places within the olivine crystal during the prograde transformation, rather than being associated with a post-shock, retrograde process. Here, we present further evidence to support our suggestion that the development of stacking faults in the spinel phase is indeed related to the mechanism of the spinel to β -phase transformation.

The Peace River meteorite¹² (classified as an L6 olivine-hypersthene chondrite) contains shock-produced veins, which pervade the body of the meteorite in much the same way as the veins found in other shocked chondrites (for example, Tenham, Coorara, Catherwood and Coolamon)^{13–16}. In fragments found within these veins, the constituent olivines and hypersthene have been transformed to their high-density polymorphs, ringwoodite and the garnet-structure majorite respectively. In general, these 'high-pressure phases' have remained, quenched-in after an extraterrestrially produced shock event. The vein studied in the Peace River meteorite differs significantly, however, from those described in other meteorites, in that it contains significant quantities of the β -phase polymorph of $(\text{Mg, Fe})_2\text{SiO}_4$. The β -phase occurs as microcrystalline aggregates which pseudomorph previously existing olivine fragments within the vein. The fragments rarely exceed 0.5 mm in diameter. Electron microprobe analyses of β -phase aggregates yield an average, ideal formula for the observed β -phase of $(\text{Mg}_{1.4}\text{Fe}_{0.5}\text{Mn}_{0.1})\text{SiO}_4$. Representative analyses of a β -phase aggregate and an untransformed olivine are given in Table 1; within experimental limits, no compositional differences between the $(\text{Mg,Fe})_2\text{SiO}_4$ polymorphs could be detected. Least-squares refinement of X-ray powder data obtained from separated β -phase fragments yielded orthorhombic cell parameters of $a = 5.70(2)$, $b = 11.51(7)$ and $c = 8.24(4)$. The calculated cell volume for the observed β -phase is $541(3)\text{\AA}^3$, and with $z = 8$, the calculated density on the basis of the idealized formula is 3.84 g cm^{-3} .

Transmission electron microscopy (TEM) shows that relatively large fragments within the shock-vein are composed of granular aggregates of well formed, micrometre-size β -phase grains, with intergranular boundaries usually meeting at triple points (often with interfacial angles close to 120°). Grains of β -phase are also found replacing grains of ringwoodite (Fig. 1). The spinel grains invariably carry a high density of $a/4[110]$ (110) stacking faults (identical to those found in the Tenham meteorite^{5,6}). There is an occasional slight ($\pm 2^\circ$) misalignment of the β -phase lattice relative to the spinel lattice, but electron diffraction confirms that the phases are essentially orientated such that $[010]_\beta \parallel [110]_\gamma$ and $[001]_\beta \parallel [001]_\gamma$. The phases share, therefore a continuous, near cubic-close-packed oxygen sub-lattice.

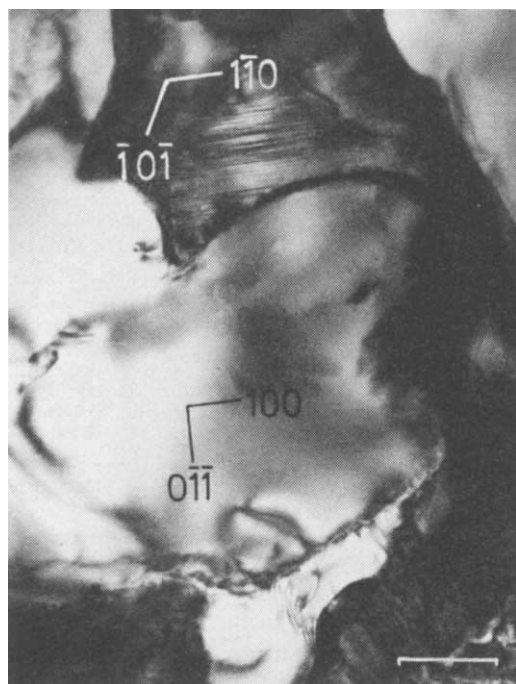


Fig. 1 Electron micrograph of a fault-free β -phase grain, topotactically replacing faulted ringwoodite. The two phases are related such that $[010]_{\beta} \parallel [110]_{\gamma}$ and $[001]_{\beta} \parallel [001]_{\gamma}$. A subgrain boundary separates the two phases, but other examples were found which had coherent interfaces. Scale bar, 0.2 μm .

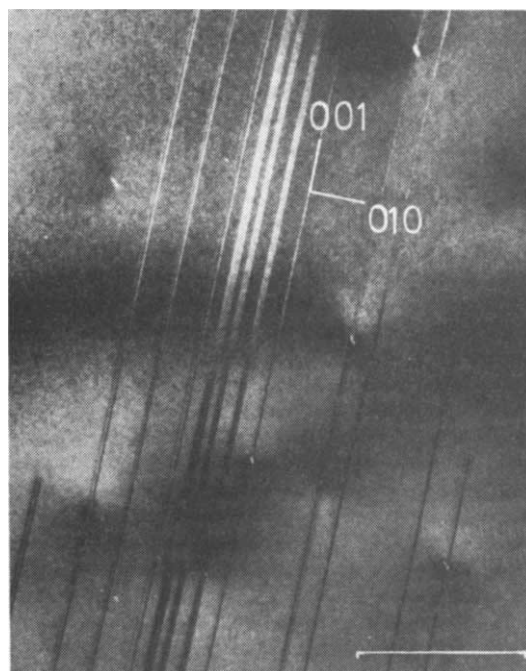


Fig. 2 Residual faults found in a β -phase grain. The faults are of the type $1/4[010]$ (010), and give rise to the spinel cation distribution within the fault plane. Faults are terminated or displaced at dislocations. Scale bar, 0.1 μm .

The textures observed are interpreted as resulting from post-shock inversion of the spinel polymorph to a lower density phase, in response to a reduction in pressure. The proposed mechanism for the observed transformation of spinel to the β -phase involves two stages. The process initially is a cooperative or martensitic one, and probably proceeds by the propagation of low-energy¹⁷, Shockley partial dislocations of the type $1/4[1\bar{1}2]$ (110) through the spinel, to produce the observed stacking faults on {110} planes. These faults form on all {110} sets of planes in the spinel, however, and prevent the simple cooperative formation of a single, β -phase grain from the host spinel (a process which would require faults on only one set of {110} planes). As the density of faults increases, a critical volume of β -phase is produced locally, which then grows and topotactically replaces the host. As the β -phase region grows, the faults in the spinel are swept before the advancing interface (Fig. 1) to produce a relatively fault-free grain of β -phase. Occasionally, relict faults are found in the β -phase grains (Fig. 2), which high resolution electron microscopy has shown to be of the type $1/4[010]$ (010) (expressed in terms of the β -phase lattice), and which are therefore complementary to those found in

spinel, that is locally retaining the spinel cation distribution within the β -phase matrix. The faults are often complex in detail and are frequently terminated at dislocations within the grains. They are interpreted as being residual features preserved as a result of the relatively rapid, post-shock growth process, and are the subject of continuing research.

The spinel to β -phase inversion is expected to occur^{7,18} within the mantle, at a depth close to 400 km, as a result of the peritectoidal relationship between the olivine, spinel and β -phase polymorphs of magnesium-rich members of the $(\text{Mg,Fe})_2\text{SiO}_4$ system. However, considering the different thermal regimes involved, it is difficult to assess whether the transformation of spinel to β -phase occurs in the mantle by the same mechanism as the one which has been inferred from the textures in the Peace River meteorite. A mechanism of this type occurring in the mantle would have as the rate-controlling step the diffusion of atoms across the β -phase/spinel interface. The initial cooperative or martensitic stage of the mechanism would, however, particularly weaken the spinel with respect to shearing stresses on {110}. Suitable shears may allow the transformation to proceed to completion by a single, military process, and in this case, the rate-controlling step would be the formation and glide of partial dislocations through the spinel lattice. On the basis of these suggested mechanisms, it may be inferred, therefore, that the overall kinetics of the spinel to β -phase transformation will be particularly sensitive to shear. In addition, the resistance of spinel bearing mantle to shear stresses would be significantly reduced during the transformation to the β -phase, and would affect mantle rheology.

We acknowledge the help and advice of Dr S. O. Agrell, and the NERC for electron microscope facilities. G.P.D. also acknowledges the receipt of Research Fellowships from the NERC and from Clare College, Cambridge.

Received 22 December 1981; accepted 26 February 1982.

1. Bernal, J. D. *Observatory* **59**, 268 (1936).
2. Ringwood, A. E. & Major, A. *Phys. Earth planet. Inter.* **3**, 89 (1970).
3. O'Keeffe, M. & Hyde, B. G. *Nature* **293**, 727 (1981).
4. Sung, C.-M. and Burns, R. G. *Phys. Chem. Miner.* **2**, 177 (1978).
5. Putnis, A. & Price, G. D. *Nature* **280**, 217 (1979).
6. Price, G. D., Putnis, A. & Agrell, S. O. *Contr. Miner. Petrol.* **71**, 211 (1979).

Table 1 Electron microprobe analyses of β -phase aggregates and shocked olivine crystals from the Peace River meteorite

	β -phase (wt%)	Olivine (wt%)
SiO_2	38.37	38.16
Al_2O_3	0.00	0.00
Cr_2O_3	0.00	0.00
MgO	38.31	38.38
CaO	0.05	0.00
MnO	0.31	0.40
FeO	23.10	23.57
CoO	0.20	0.00
NiO	0.11	0.05
Total	100.45	100.56

7. Poirer, J. P. *Phys. Earth planet. Inter.* **26**, 176 (1981).
8. Madon, M. & Poirer, J. P. *Science* **207**, 66 (1980).
9. Veyssiere, P., Rabier, J. & Grilhe, J. *Phys. Stat. Sol. A* **31**, 605 (1975).
10. Morimoto, N., Akimoto, S., Koto, S. & Tokohami, M. *Phys. Earth planet. Inter.* **3**, 161 (1970).
11. Moore, P. B. & Smith, J. V. *Phys. Earth planet. Inter.* **3**, 166 (1970).
12. Folinsbee, R. E. & Bayrock, L. A. *J. R. astr. Soc. Canada* **58**, 109 (1964).
13. Binns, R. A. *Phys. Earth planet. Inter.* **3**, 156 (1970).
14. Smith, J. V. & Mason, B. *Science* **168**, 832 (1970).
15. Coleman, L. C. *Can. Miner.* **15**, 97 (1977).
16. Steele, I. M. & Smith, J. V. *Lunar planet. Sci.* **9**, 1101 (1978).
17. Van der Biest, O. & Thomas, G. *Phys. Stat. Sol. A* **24**, 65 (1974).
18. Akimoto, S., Matsui, Y. & Syono, Y. in *Physics and Chemistry of Minerals and Rocks* (ed. Strens, R. G. J.) (Wiley, London, 1974).

Winchite and the actinolite–glaucophane miscibility gap

W. V. Maresch, O. Medenbach & A. Rudolph*

Institut für Mineralogie, Ruhr-Universität Bochum,
D-4360 Bochum, FRG

Leake *et al.*¹ have recently emphasized the apparent scarcity of winchite *s.l.*, $\text{NaCa}(\text{Mg}, \text{Fe}^{2+})_4(\text{Al}, \text{Fe}^{3+}) [\text{Si}_6\text{O}_{22}/(\text{OH})_2]$ (ref. 2), an amphibole intermediate in composition between the glaucophane–riebeckite series of high-pressure blueschists and the tremolite–actinolite series of lower-pressure greenschists. The rarity of winchite is considered consistent with the existence of a miscibility gap^{3,4} between these two amphibole series at the metamorphic *P–T* conditions normally attained in nature. Here we present optical and X-ray properties of Mg- and Al-rich winchite from Venezuela and draw attention to several other recent winchite analyses. In fact, as Schliestedt's⁵ new data on co-existing sodic and calcic amphiboles in blueschists show, only the present Mg–Al-rich winchite definitely lies within the above miscibility gap.

Winchite was identified in an amphibole–eclogite found in talus near El Valle del Espiritu Santo, Margarita Island, Venezuela (63°53'10" W, 10°58'55" N; Venezuela grid ref. no. 40322142). Field relationships of similar occurrences on Margarita indicate that the eclogite is a metabasalt pod in quartz–mica schist of epidote–amphibolite facies grade. The bulk-rock chemistry of the eclogite indicates a tholeiite of MORB affinity⁶. Winchite is found with omphacite, garnet, rutile and quartz, as well as with minor epidote, albite, opaque oxide, sphene and apatite. Thin actinolitic rims overgrow some of the winchite grains.

Electron microprobe analyses (Table 1) indicate that the amphiboles are winchites grading to aluminowinchites ($\text{Al} > 1.0$)². Figure 1 shows that the departure from ideal end-member winchite composition is due mainly to a minor amount of edenite substitution, $(\text{Na}, \text{K}) + \text{Al} \rightarrow \text{Si}$, giving A-site occupancy totals of 0.02–0.18 atoms per half-unit-cell formula, and more extensive Tschermak's substitution, $(\text{Al}, \text{Fe}^{3+}) + \text{Al} \rightarrow (\text{Mg}, \text{Fe}^{2+}) + \text{Si}$, leading to Si totals of 7.56–7.78. Figure 1 also gives published data of samples that classify as winchite, although the samples are markedly more iron-rich and would be classified as ferro- or ferri-winchite². All analyses are from blueschist terranes. Clearly, winchites approaching the ideal end-member are not uncommon in nature.

We have determined the optical and X-ray properties of the Venezuelan winchite (Table 2) on a single cleavage fragment $\sim 50 \times 50 \times 200 \mu\text{m}$ in size. The optical properties were corroborated on several other fragments and are considered to be representative. Indices of refraction and $2V$ were measured using the microrefractometer spindle stage of Medenbach⁷, while X-ray data were collected on a four-circle automatic diffractometer. A scan of 278 reflections between 0 and 35° MoK_α revealed no reflections violating monoclinic *C* symmetry.

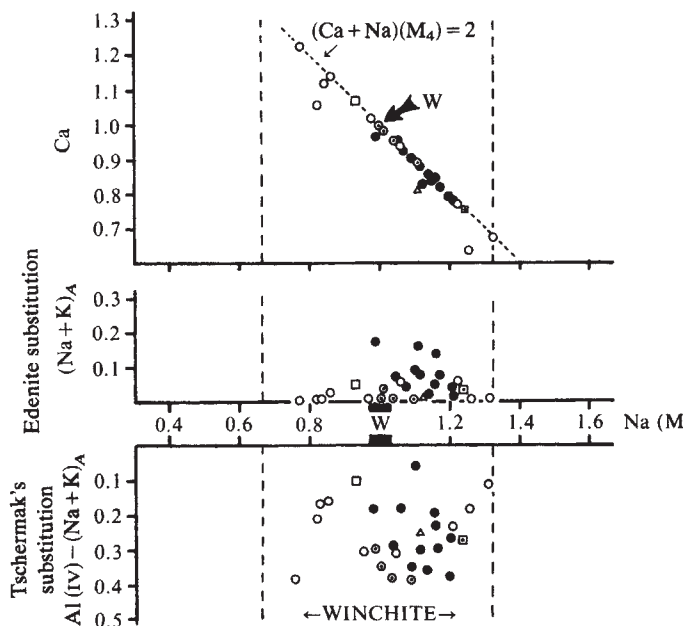


Fig. 1 Plot of $\text{Na}(\text{M}_4)$ against edenite substitution, $(\text{Na} + \text{K})_A$, Tschermak's substitution, $\text{Al}(\text{IV}) - (\text{Na} + \text{K})_A$, and Ca all referred to half-unit-cell formula. Data sources: ●, present study; ○, ref. 5; □, ref. 14; △, ref. 8; ◇, ref. 15; ◇, ref. 9. $\text{Fe}^{3+}/\text{Fe}^{2+}$ ratios for all analyses are calculated as for Table 1. W, ideal end-member winchite.

Table 1 Representative microprobe analyses of Venezuelan winchite

Analysis 1.1 (aluminowinchite)				Analysis 8.2 (winchite)			
SiO_2	55.00	Si	7.70	SiO_2	54.41	Si	7.63
TiO_2	—	Al	0.30	TiO_2	0.20	Al	0.37
Al_2O_3	8.12	Total	8.00	Al_2O_3	7.00	Total	8.00
Fe_2O_3	3.53			Fe_2O_3	4.72		
FeO	7.10	Al	1.04	FeO	6.09	Al	0.79
MnO	0.04	Ti	—	MnO	0.04	Ti	0.02
MgO	13.18	Fe^{3+}	0.37	MgO	14.25	Fe^{3+}	0.50
CaO	5.52	Fe^{2+}	0.83	CaO	6.32	Fe^{2+}	0.71
Na_2O	4.45	Mn	0.01	Na_2O	4.01	Mn	0.01
K_2O	0.12	Mg	2.75	K_2O	0.16	Mg	2.97
Total	97.06	Total	5.00	Total	97.20	Total	5.00
		Ca	0.83			Ca	0.95
		Na	1.21			Na	1.09
		K	0.02			K	0.03
		Total	2.06			Total	2.07

The $\text{Fe}^{3+}/\text{Fe}^{2+}$ ratio was calculated by normalizing tetrahedral and octahedral occupants to 13 at constant 23 oxygens. This procedure is acceptable for sodic–calcic amphiboles^{5,10}.

There is no evidence of exsolution. Table 2 shows that the optical and unit cell properties of Venezuelan winchite lie between those of tremolite and glaucophane, as expected.

The occurrence of winchite amphibole has been interpreted by Brown⁸ and Green and Spiller⁹ to indicate closure of the tremolite/actinolite–glaucophane/riebeckite miscibility gap. Recent data from Schliestedt⁵ on coexisting amphibole pairs from Sifnos, Greece, indicate, however, that the multi-dimensional nature of the miscibility gap must be considered. Two projections of the latter, based on Schliestedt's data, are given in Fig. 2. Although the exact shape of the miscibility gap in $\text{Na}/\text{Na} + \text{Ca} - \text{Fe}^{2+}/\text{Fe}^{2+} + \text{Mg} - \text{Fe}^{3+}/\text{Fe}^{3+} + \text{Al}$ space is unknown, simple geometrical reasoning suggests that only the Venezuelan winchite appears to lie in the gap with any certainty. Thus only magnesian, aluminowinchites indicate *P–T* conditions exceeding the critical temperature of the gap. Iron-rich

* Present address: Im Bungert 12, 4300 Essen, FRG.

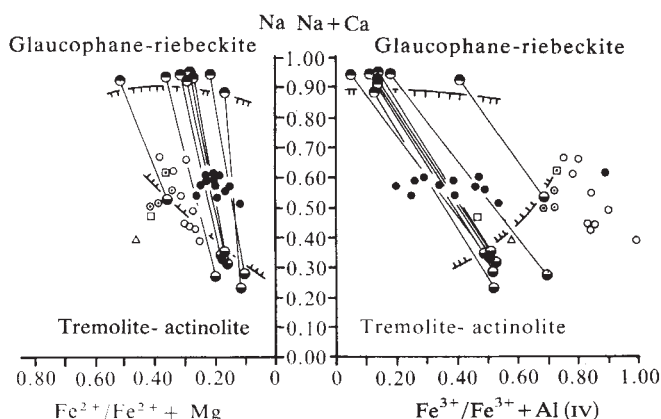


Fig. 2 Location of actinolite/tremolite–glaucophane/riebeckite miscibility gap in blueschists after data on coexisting amphiboles from Schliestedt⁵ (●) compared with winchite compositions from Venezuela and the literature. The hatched boundary represents approximate minimum solvus location. Other symbols as for Fig. 1.

Table 2 Optical properties and lattice parameters of Venezuelan winchite compared with tremolite^{11,12} and glaucophane¹³ chosen to reflect similar Fe^{2+}/Mg and Fe^{3+}/Al ratios

	Tremolite	Winchite	Glaucophane
<i>a</i>	9.818 (5)	9.71 (1)	9.595 (2)
<i>b</i>	18.047 (8)	17.93 (2)	17.798 (3)
<i>c</i>	5.275 (3)	5.283 (6)	5.307 (1)
β	104.65 (5)	104.3 (1)	103.66 (1)
<i>V</i>	904.2 (6)	891 (2)	880.6
<i>n_x</i>	1.622	1.6294 (5)	1.626
	Pale yellow-green	Colourless	Nearly colourless
<i>n_y</i>	1.633	1.6428 (5)	1.644
	Pale green	Light blue-violet	Light purple
<i>n_z</i>	1.644	1.6496 (5)	1.646
	Pale blue-green	Light blue	Light blue
Δ	0.022	0.020	0.020
$2V_x$	80–85°	64 ± 2°	38.1°
<i>Zc</i>	14°	16°	6°

winchite can be stable in normal blueschist-facies conditions where magnesian, aluminowinchite is not. Thus the latter should not be found in typical blueschists.

The upper temperature stability limit for winchite is unknown. For most common metabasaltic compositions metamorphosed in intermediate- to high-pressure conditions, however, increasing Tschermak's substitution with increasing temperature in the amphibole of the basic assemblage will lead to more Si-deficient amphibole compositions such as barroisite or hornblende. Thus the common range of *P*–*T* stability of magnesian, aluminowinchite must be severely restricted. For the Venezuelan, winchite-bearing amphibole eclogite, Rudolph⁶ has estimated equilibration conditions of at least 10 kbar and 600 ± 30 °C for the pyroxene–garnet pair, but it is not clear whether winchite growth was approximately coeval or occurred significantly later in the metamorphic cycle.

We thank W. Gebert for supplying the X-ray data, and W. Schreyer and R. Grapes for helpful discussions. The research was partly supported by grant Ma 689/1 of the Deutsche Forschungsgemeinschaft, Bonn, to W.V.M.

Received 4 January; accepted 5 March 1982.

1. Leake, B. E., Farrow, C. M. & Nayak, V. K. *Am. Miner.* **66**, 625–631 (1981).
2. Leake, B. E. *Can. Miner.* **16**, 501–520 (1978).
3. Iwasaki, M. *J. Fac. Sci. Tokyo Univ. Sec. II* **15**, 1–90 (1963).
4. Coleman, R. G. & Papike, J. J. *J. Petrol.* **9**, 105–122 (1968).
5. Schliestedt, M. thesis, Braunschweig Univ. (1980).
6. Rudolph, A. thesis, Ruhr-Univ. Bochum (1981).
7. Medenbach, O. *Fortschr. Miner.* **58**(1), 90–91 (1980).
8. Brown, E. H. *Bull. geol. Soc. Am.* **85**, 333–344 (1974).

9. Green, T. H. & Spiller, A. P. *Am. Miner.* **62**, 164–166 (1977).
10. Maresch, W. V. & Abraham, K. J. *Petrol.* **22**, 337–361 (1981).
11. Tröger, W. E. *Optische Bestimmung der Gesteinsbildenden Minerale* Pt 1 (Schweizerbart, Stuttgart, 1971).
12. Borg, I. Y. & Smith, D. K. *Mem. geol. Soc. Am.* **122**, 1–896 (1969).
13. Borg, I. Y. *Contr. Miner. Petrol.* **15**, 67–92 (1967).
14. Hermes, O. D. *J. Petrol.* **14**, 1–32 (1973).
15. Katagas, C. *Contr. Miner. Petrol.* **46**, 257–264 (1974).

History of the continental lithosphere recorded by ultramafic xenoliths

Claude J. Allègre, Nobu Shimizu* & Dominique Rousseau

Laboratoire de Géochimie et Cosmochimie, Institut de Physique du Globe de Paris, Université de Paris 6, 75230 Paris Cedex 05, France

Since the development of the plate tectonic theory, the formation, structure and properties of the continental lithosphere have been widely debated^{1–3}. By analogy with petrological models for the oceanic lithosphere generated at ridges^{4–6}, the continental lithosphere has been interpreted in terms of a comprehensive mechanical and petrogenetic unit by considering, for example, that the continental crust was differentiated through a process such as zone melting, the lithospheric mantle being the residue of this extraction process^{7,8}. However, geological observations show that formation of the continental crust was quite complex, including sedimentation, metamorphism and orogenesis, and not simply direct magmatic segregation. Underneath its crust, the continental lithosphere, mainly comprises peridotites and in this respect, it resembles the oceanic lithosphere. This similarity, which is corroborated by seismic data, does not, however, mean that a chemical coupling exists between the subcontinental lithospheric mantle and continental crust in a similar way as for the oceanic lithosphere. Although magma generation and eruption through the continental lithosphere contribute to the growth of the continental crust and to the depletion of the subcontinental lithospheric mantle in incompatible elements, no simple chemical coupling should exist if the continental lithosphere were created by thermal accretion. The petrology and geochemistry of the continental lithosphere can also be studied using xenoliths brought up to the surface by kimberlitic or alkalic volcanics. This technique attracted much attention because of the determination of pyroxene geotherms⁹ and the claim that it permits the description of the palaeogeothermal structure of continents and oceans^{10,11}. Although some of the original conclusions seem premature¹² the basic idea is sound. We have now used radiogenic tracers, the ⁸⁷Rb–⁸⁷Sr system, to study the creation of the continental lithosphere. We present here results on garnet lherzolite xenoliths from southern African kimberlites and discuss their geochemical and geodynamical implications.

The major and trace element chemistry of the constituent minerals of most of the samples studied here (see Table 1) has been described elsewhere¹³. The olivine-bearing xenoliths in kimberlites can be ascribed to two textural groups: granular peridotites and sheared lherzolites. Whereas the sheared samples have a 'pristine-mantle-type' chemical composition, the granular ones show a very complex chemistry, depleted in the basaltic components while enriched in some trace elements. Clear evidence for mantle metasomatism has also been reported¹³ for some granular samples.

To study the isotopic characteristics of xenoliths, the minerals enriched in the elements of interest and which are insensitive to secondary reaction processes must be separated. Thus, we separated diopside and garnet as they are the main carriers of

* Present address: Center of Geoalchemy, Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

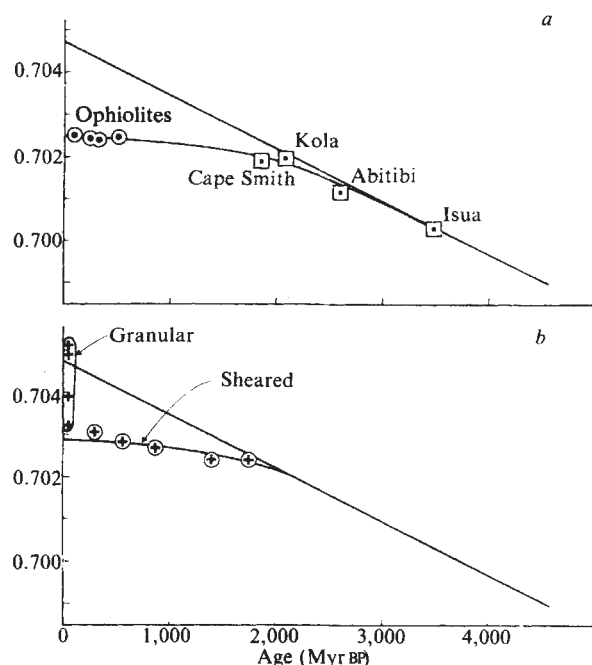


Fig. 1 *a*, ^{87}Rb - ^{87}Sr diagram showing the thermal evolution of a garnet peridotite under rehomogenization; *b*, ^{87}Rb - ^{87}Sr diagram for sheared xenoliths from kimberlites showing the localization of several garnet-diopside isotopic tie-lines.

Sr and Rb in these rocks. Separation was done by hand-picking, the purity was checked under a microscope and the separates were washed in high-purity acetone and water to eliminate any surface contamination. The Sr isotopes were analysed using the technique described in ref. 14; the blanks and the results on the standards are given in Table 2. Rb and Sr concentrations were obtained by an isotope dilution technique.

The analytical results summarized in Table 3 show that diopsides have very low Rb/Sr ratios and thus tend to fossilize the $^{87}\text{Sr}/^{86}\text{Sr}$ initial ratio of a mineral assemblage. The observed variation in strontium isotopic composition among diopsides (0.70245–0.70512) clearly demonstrates the heterogeneity of the initial ratio for the continental lithospheric mantle at the time of the kimberlite eruptions.

Such an heterogeneity may reflect either an intrinsic isotopic heterogeneity of the mantle rocks if the assemblages are in internal equilibrium, or different episodes of formation of the continental lithosphere corresponding to the different initial ratios. In an attempt to answer this question we also analysed the garnets, which are the second largest reservoir of Sr after diopside.

The most remarkable result is that diopside and garnet in a given granular sample have very similar $^{87}\text{Sr}/^{86}\text{Sr}$ ratios, whereas the sheared lherzolites yield very different Sr isotope

ratios for these minerals. In other words, considering the ^{87}Sr growth since the time of kimberlite eruption, diopside and garnet are close to internal isotopic equilibrium for the granular xenoliths while they are in isotopic disequilibrium for the sheared samples. This behaviour is most surprising in the light of the geothermometry data, as the granular xenoliths typically yield a temperature around 1,000 °C compared with 1,300 °C for the sheared samples. How can rocks with large crystals such as the granular samples be isotopically homogenized at 1,000 °C whereas higher temperature facies with a finer grain-size are in isotopic disequilibrium?

If the rate of re-equilibration (presumably controlled by the diffusion rate of elements in minerals) were much greater for the major elements than for Rb and Sr, the Sr isotopic disequilibrium could be preserved while the major elements would be completely re-equilibrated. Diffusion of major and trace elements in silicate minerals is still poorly understood. Recent

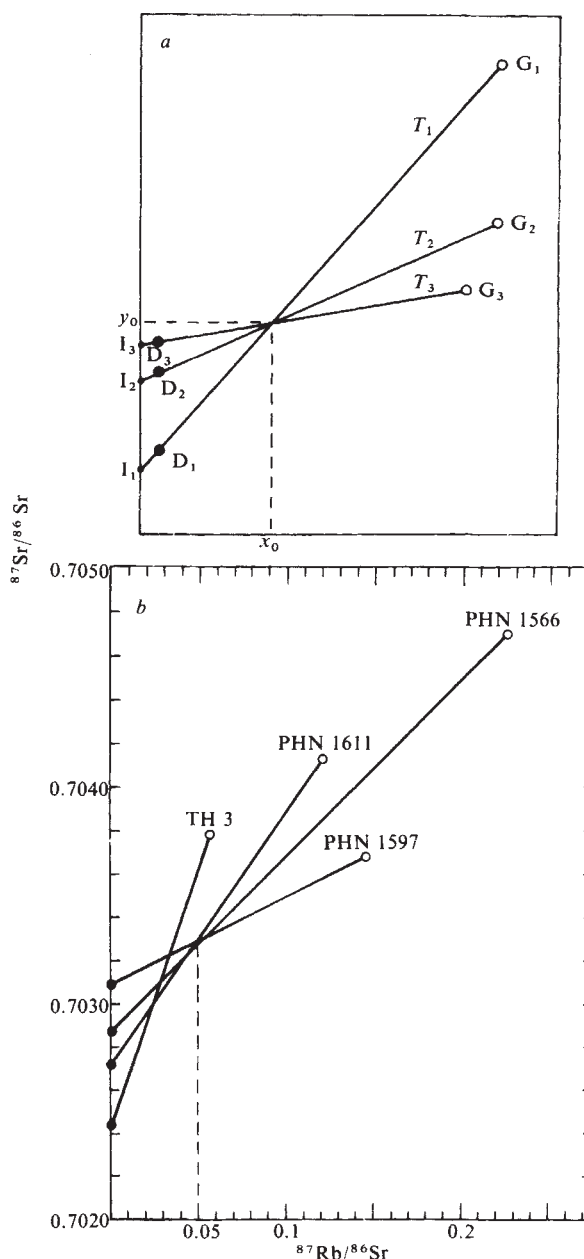


Fig. 2 *a*, ($^{87}\text{Sr}/^{86}\text{Sr}$, time) isotopic evolution diagram for the sub-oceanic mantle, after ref. 19; *b*, (apparent age, $^{87}\text{Sr}/^{86}\text{Sr}$ initial ratio) diagram for peridotite xenoliths; sheared nodules Thaba Putsoa.

Table 1 List of samples

No.	Locality	Textural type	Chemical type
W 397	Premier Mine	Porphyroclastic	I
PHN 1611	Thaba Putsoa	Fluidal mosaic	I
PHN 1597	Thaba Putsoa	Mosaic	I
PHN 1566	Thaba Putsoa	Mosaic	I
THA 3	Thaba Putsoa	Porphyroclastic	I
PHN 1569	Thaba Putsoa	Coarse granular	III
AJE 25	Bultfontein	Coarse granular	II
MAT 7	Matsoku	Coarse granular	II
MAT 10	Matsoku	Coarse granular	(Spinel lherzolite)

After ref. 13.

work¹⁵ on Ca self-diffusion suggests a diffusion coefficient varying from $7.0 \times 10^{-12} \text{ cm}^2 \text{ s}^{-1}$ to $2.2 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$ between 1,300 and 1,200 °C, respectively, with an upper limit of $2 \times 10^{-12} \text{ cm}^2 \text{ s}^{-1}$. For Fe-Mg interdiffusion in Ca-rich pyroxene, only Sanford and Huebner's¹⁶ calculated minimum value of $4 \times 10 \text{ cm}^2 \text{ s}^{-1}$ at 1,050 °C is available. Sneeringer and Hart¹⁷ obtained a diffusion coefficient for Sr in diopside of $2.7 \times 10^{-12} \text{ cm}^2 \text{ s}^{-1}$ in the range 1,250–1,300 °C. None of these values is very well established, but they seem to suggest that there are no significant differences between the rate of re-equilibration of major elements and Sr. One may thus speculate that, for a given mineral couple, Sr isotopes are equilibrated only if the major elements have also reached equilibrium.

Therefore, we conclude that pyroxene geotherms or similar features have no simple interpretation. More probably, the (*P*, *T*) determination corresponds to fossil conditions, but different rocks have recorded these conditions at different times.

If the minerals in the sheared nodules form an internal isochron of age *T*₁ relative to a parent-daughter system, the *P*-*T* conditions calculated for the minerals correspond to the time *T*₁. If this were indeed the case, it would call for a process in which high-temperature equilibrium was frozen and preserved for a long time. Hofmann and Hart¹⁸ stated that "subcontinental lithosphere is essentially protected against high temperatures and reworking by virtue of the overlying crust" and the xenoliths could come from mantle levels where the temperature has long been below the closure temperature for diffusion of the major elements as well as Sr.

Figure 1a is an isochron diagram showing the diopside-garnet joins for sheared xenoliths from the Thaba-Putsoa pipe. Note that these tie lines intersect each other almost at a unique point. Furthermore, the steeper the slope of the tie-lines, the lower the ⁸⁷Sr/⁸⁶Sr ratios for the diopside. We consider this as a strong evidence against impurity effects, particularly on garnet separates, as we expect that random impurity effects would yield random intersections of the tie lines. Thus, we would closely approximate the ideal situation depicted in Fig. 1b in which the system is re-equilibrated isochemically at different times (*T*₁, *T*₂, and so on). The fact that diopside and garnet practically account for the total content in Rb and Sr of the garnet lherzolites supports our hypothesis. An important geochemical implication is that the sheared xenoliths from Thaba Putsoa represent various portions of a mantle system with a common Rb/Sr ratio, the value of which is quite low. The fact that the sheared-lherzolite mineral phase (and thus the whole rocks) show a remarkable coherence relative to some trace elements¹³ is consistent with our speculation.

The situation illustrated in Fig. 1b may be readily converted into another linear relationship by plotting slopes against intercepts. Because Fig. 1 is an isochron diagram, the slopes rep-

Table 3 Analytical results for kimberlite nodules

	⁸⁷ Sr/ ⁸⁶ Sr	⁸⁷ Rb/ ⁸⁶ Sr	Rb (p.p.m.)	Sr (p.p.m.)
Granular nodules				
MAT 7				
Diopside	0.70323 ± 10	0.00087	0.081	270
Enstatite	0.70380 ± 12	0.381	1.40	10.64
Garnet	0.70408 ± 16	1.53	1.10	2.08
MAT 10				
Diopside	0.70496 ± 10	0.00042	0.013	91.5
Olivine	0.70799 ± 10	0.00058	0.014	6.95
PHN 1569				
Diopside	0.70398 ± 6	0.00017	0.031	514
Enstatite	0.70379 ± 12	0.0363	0.104	8.33
AJE 25				
Diopside	0.70512 ± 10	0.00044	0.032	209
Enstatite	0.70494 ± 7	0.234	0.440	5.50
Garnet	0.7056 ± 2	0.252	0.105	1.21
Sheared nodules				
THA 3				
Diopside	0.70243 ± 6	0.00056	0.0013	83.3
Enstatite	0.7051 ± 4	0.0103	0.020	5.61
Garnet	0.70379 ± 15	0.056	0.0427	2.20
W397				
Diopside	0.70245 ± 7	0.00097	0.034	98.3
Garnet	0.7253 ± 2	1.10	0.844	2.21
RT	0.71772 ± 9	0.784	2.79	10.3
PHN 1611				
Diopside	0.70272 ± 4	0.0018	0.074	116
Enstatite	0.70397 ± 10	0.460	0.946	5.94
Garnet	0.70413 ± 15	0.115	0.426	10.7
RT	0.7035 ± 1	0.562	7.82	40.3
PHN 1566				
Diopside	0.70285 ± 6	0.00034	0.011	94.3
Garnet	0.7047 ± 3	0.224	0.182	2.35
Enstatite	0.70489 ± 14	0.113	0.134	3.44
PHN 1597				
Diopside	0.70309 ± 9	0.0012	0.035	82.6
Garnet	0.70367 ± 6	0.143	0.076	1.54
Olivine	0.70383 ± 15	0.379	0.611	4.48

resent times and the intercepts, the initial ⁸⁷Sr/⁸⁶Sr ratios and hence the Sr isotopic characteristics of this particular mantle system can be further examined in terms of its Sr isotopic evolution. Figure 2 shows a good linear fit, not only for the four sheared xenoliths from Thaba Putsoa but also for a sample from the Premier Mine pipe (1,400-Myr old kimberlite).

Compared with the bulk earth Sr isotope evolution curve, the sheared xenoliths define a Sr isotopic evolution for a depleted mantle; this evolution trend is very similar to that for the oceanic lithosphere¹⁹. If we calculate the ⁸⁷Rb/⁸⁶Sr ratio for the uncontaminated whole rocks, through the isotopic evolution diagram (Fig. 1b), we get values around 0.05, which undoubtedly indicates a depletion since the bulk earth value is 0.09.

The Sr isotopic characteristics summarized above may be interpreted in terms of continuous or episodic thermal accretion of the continental lithosphere at its base. When an old shield continuously accretes this lithosphere the earliest fraction producing the uppermost lithospheric mantle is derived from older mantle which, until that time, has been undepleted; the youngest fraction producing the lower part of the continental lithosphere shows the reverse trend.

In contrast, we can admit that continental lithosphere was mostly created just after the local orogenic processes which built the local continental crust. The subsequent cooling of such episodic accreting lithosphere is not uniform with depth. Therefore, the isotopic freezing occurs at different times for different depths, creating the observed isotopic zonation.

In either model, we must consider independently the granular xenolith data. These samples are depleted in major elements but some are quite undepleted in terms of their ⁸⁷Sr/⁸⁶Sr ratios. The characteristics of these xenoliths may be explained through a two-stage model. At their depth level, melting occurred,

Table 2 Isotopic analysis results

Blanks			
Date	K (ng)	Rb (pg)	Sr (pg)
26 April 1978	0.46	26	58
22 May 1978	0.83	59	92
9 June 1978	—	13	45
16 October 1978	37	43	117
Standards			
	$^{87}\text{Sr}/^{86}\text{Sr} \pm \frac{2\sigma}{\sqrt{n}}$		
12 March 1978	0.71016 ± 0.00006		
3 April 1978	0.71018 ± 0.00007		
14 April 1978	0.71015 ± 0.00006		
16 May 1978	0.71016 ± 0.00006		
19 July 1978	0.71010 ± 0.00009		
16 October 1978	0.71007 ± 0.00005		
23 October 1978	0.71012 ± 0.00004		

creating some continental basalts such as the Karroo basalts. The granular facies would thus represent residues after basalt extraction events which are indeed quite recent. However, some of the granular samples correspond to higher level and they have a $^{87}\text{Sr}/^{86}\text{Sr}$ signature quite undepleted, others correspond to a lower level and are more depleted. Such volcanism extracted basalts, leaving residual material such as the granular-facies rocks, yield fluids which induced mantle metasomatism. The combination of the two phenomena created the complexities in the chemistry and isotopic characteristics of the continental lithosphere.

We thank J. C. Mercier, S. R. Hart, and B. Dupré for helpful discussions. This is I.P.G.P. contribution NS/571.

Received 12 October 1981; accepted 12 February 1982.

- McDonald, G. F. *Rev. Geophys.* **587**–665 (1963).
- Jordan, T. H. *Nature* **274**, 544–548 (1978).
- Okal, E. A. & Anderson, D. L. *Geophys. Res. Lett.* **2**, 313–316 (1975).
- Oxburg, E. R. & Turcotte, D. L. *J. geophys. Res.* **73**, 2643–2661 (1968).
- Bottinga, Y. & Allegre, C. J. *Tectonophysics* **18**, 1–17 (1973).
- Slater, J. G. & Francheteau, J. *Geophys. J. R. astr. Soc.* **20**, 509–550 (1970).
- Kono, Y. & Amano, M. *Geophys. J. R. astr. Soc.* **54**, 405–416 (1978).
- Wells, P. R. A. *Earth planet. Sci. Lett.* **46**, 253–265 (1980).
- Boyd, F. R. & Nixon, P. H. *Yb. Carnegie Instn Washington* **72**, 431–445 (1973); *Phys. Chem. Earth* **9**, 431–454 (1975).
- MacGregor, I. D. *Am. Miner.* **59**, 110–119 (1974).
- Mercier, J.-C. C. & Carter, N. L. *J. geophys. Res.* **80**, 3349–3362 (1975).
- Wilshire, H. G. & Jackson, E. D. *J. Geol.* **83**, 313–329 (1975).
- Shimizu, N. & Allegre, C. J. *Contr. Miner. Petrol.* **67**, 41–50 (1978).
- Birck, J. L. & Allegre, C. J. *Earth planet. Sci. Lett.* **20**, 266–270 (1973).
- McCallister, R. H., Brady, J. B. & Mysen, B. O. *Yb. Carnegie Instn Washington* **78**, 574 (1979).
- Sanford, R. S. & Huebner, J. F. *Lunar planet. Sci.* **10**, 1052 (1979).
- Sneeringer, M. & Hart, S. R. *EOS* **59**, 402 (1978).
- Hofmann, A. & Hart, S. R. *Earth planet. Sci. Lett.* **38**, 44–62 (1978).
- Allegre, C. J. *Tectonophysics* (in the press).

Origin of longitudinal triangular ripples on the Nova Scotian continental rise

Brian E. Tucholke

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

Since the recognition of current-produced bedforms in the deep ocean, attempts have been made to predict direction, intensity and variability of abyssal flow from the nature of sea floor irregularities (see refs 1–3). Longitudinal triangular ripples (LTRs) are intermediate-scale bedforms (<20 cm high, metres long) that are becoming recognized as important and geographically widespread predictors of bottom-current flow. I describe here recent sea floor photographs of LTR fields on the lower continental rise off Nova Scotia which show that the bedforms develop in fine-grained, cohesive sediment with an alignment parallel to the regional contours and to mean current direction. An observational model for LTR formation is proposed, wherein LTRs are deposited initially as 'tails' behind large sea floor biological mounds and are subsequently constructed and propagate in the direction of mean current by sediment transport oblique to the bedform axis. This transport is induced by short-term currents that deviate from mean current direction and cause formation of a separation bubble on one side or the other of the LTR crest.

Current-produced bedforms in the deep ocean range from mere lineations a few millimetres high to abyssal sediment waves up to 100 m high and several kilometres in wavelength⁴. Among the larger bedforms that can be photographed with conventional deep-sea cameras are LTRs ranging up to ~15 cm high, 1 m wide and several metres in length. Such features appear to align with the mean current direction, and they have been photographed at depths of 4,600–5,400 m along the lower continental rise off eastern North America^{5–7}, at the foot of the Bahama Banks⁸, on the north-east Bermuda Rise⁹, and in the

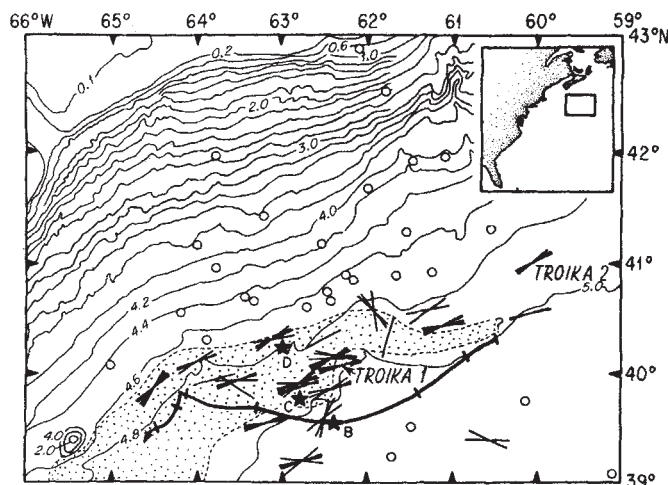


Fig. 1 Location and orientation of longitudinal triangular ripples¹⁶. Note that even NW–SE oriented LTRs at right centre parallel local bathymetric contours. Troika-1 was a 3.4-km photographic transect and Troika-2 an 8.1-km transect. ○, Photograph stations with no LTRs. Surface sediment in stippled area contains more than 10% sand. Bathymetry in corrected metres. ★, Locations of current-meter records: B, 15-day record; C, 215-day; D, 230-day; progressive vector diagram at B is tick-marked at 2-day intervals.

south-west Indian Ocean¹⁰. Recent detailed sea floor investigations for the high energy benthic boundary layer experiments (HEBBLE)¹¹ have shown that LTRs are widespread at depths of 4,700–5,100 m on the lower continental rise off Nova Scotia (Fig. 1). The bedforms have been photographed with both a Ewing-Thorndike bottom-bounce camera¹² (high-angle oblique view) and a Troika camera sled¹³ (low-angle oblique view).

The LTRs on the Nova Scotian continental rise are best developed at depths of 4,800–5,000 m. The surface sediment is a sandy silty clay (<20% sand, 30–50% silt and 30–60% clay). LTRs tend to be better developed and more frequent on the southwestern part of the continental rise where sediments also have higher sand content (Fig. 1). They generally align within 10°–15° of the strike of the bathymetric contours. Individual LTRs are up to 15 cm high, about 1–5 m long, and their spacing ranges from <1 m to several tens of metres (Fig. 2). Width ranges up to ~1 m and varies roughly in proportion to height. The ripples are mostly bilaterally symmetrical, and they are occasionally slightly sinuous. Ripple-crest height is not constant but varies by a few centimetres along the length of a ripple (Fig. 2e). Where the tails (downstream, south-west ends) of the LTRs are observed, they blend smoothly into the sea floor. In contrast, the heads of the LTRs merge smoothly with the sea floor in only about one-third of the ripples. The remaining ripples have a distinct bulge at the head, often recognized as a biologically produced mound (Fig. 2b, d). Thus it seems that the LTRs commonly, if not ubiquitously, have formed behind such obstacles to flow.

Associated with the LTRs are various other current-produced bedforms. Most common are fine lineations a few millimetres high, and crag-and-tail features up to 1 cm high and several centimetres long; these often occur on the LTRs (Fig. 2). Orientation of these smaller features is more irregular than that of larger bedforms, suggesting that there is significant short-term (days to weeks?) variability in the direction of currents that produce the small-scale bedforms. Measured currents in this area do show such variability superimposed on a mean contour-parallel flow towards the south-west¹⁴. The LTRs themselves appear to be 'vector-averaging current vanes' oriented in the direction of long-term mean flow.

The small-scale bedform patterns superimposed on the LTRs are pertinent to LTR origin. Close study of photographs shows three kinds of ripple-crest development. Least common is a sharp, symmetrical crest: more often the crest is rounded

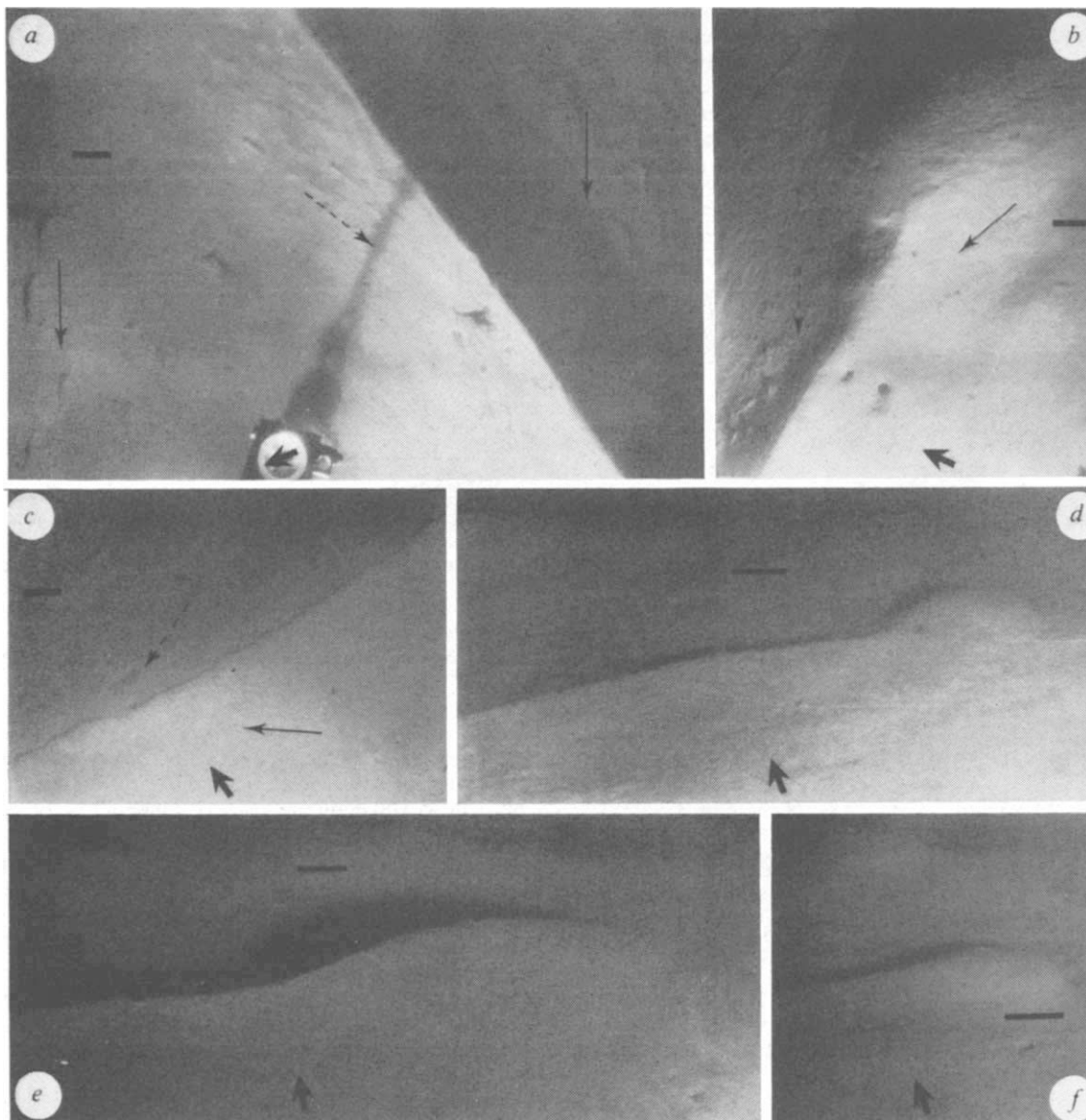


Fig. 2 Photographs of longitudinal triangular ripples. Bold arrow shows true north. *a*, LTR showing southwesterly flow within separation bubble on north flank (dashed arrow); main flow is to west (solid arrow). *b*, LTR formed behind biological mound; arrows as in *a*. *c*, LTR with cornice forming ripple crest. *d*, LTR developed behind biological mound. *e*, Mound at head of LTR, possibly of biological origin and current-modified; in *d* and *e*, changes in ripple-crest height give false appearance of sinuosity. *f*, Biological mound with well-developed tail, possibly precursor to LTR. *a*, *b*, and *c* are high-angle oblique photographs (courtesy C. D. Hollister); *c*, *d*, and *e* are low-angle oblique Troika photographs. Scale bars, 10 cm.

(Fig. 2*d*). In both cases fine current-produced lineations usually parallel the ripple crest. A third, common situation is a slightly asymmetrical crest (Fig. 2*a*, *c*). In such instances the fine-scale lineations define a short-term current trending at an angle to the crest (usually $<45^\circ$), and a small 'slip-face' or cornice can be developed on the lee edge of the crest (Fig. 2*c*). On the lee flank of the ripple, the fine-scale lineations also trend upwards towards the ripple crest. The lee-side lineations indicate a zone of flow convergence; this convergence is thought to occur in a separation bubble because farther downstream from the ripple the lineations again parallel those on the upstream flank (Figs 2*a*, 3). Laboratory experiments by Allen (ref. 15, pages 210–216) show similar formation of a separation bubble behind skewed straight steps in a flume bed. When the incident current is at $<40^\circ$ – 45° to the step (ripple) an open separation bubble is formed; it has helical internal flow (Fig. 3) that can extend beyond the tail of the ripple. A closed bubble, having internal circulation but lacking the extended helical flow, forms at greater angles of incidence (45° – 135°).

Based on these observations, I propose that longitudinal triangular ripples originate as a consequence of four factors: (1) presence of a large (~ 10 – 20 cm high) seabed obstacle to flow; (2) bedload transport and/or deposition from suspended load; (3) presence of bottom currents with a significant mean flow; and (4) definite variation in direction of flow about the mean. In this model an obstacle to flow, such as a biologically produced mound, creates a leeward zone of flow convergence (wake); a tail oriented in the direction of mean flow is deposited in this wake. Common observation in beach and dune areas shows that beneath directionally uniform flows such a tail reaches only a finite size determined by the length of the wake behind the obstacle. However, where significant variations in flow direction occur, I suggest that a separation bubble is formed on one side or the other of the initial tail in response to the cross-bedform flow (Figs 2*a*, 3). The tail (ripple) thus is further constructed by bedload and/or suspended load transport towards the bedform crest on both sides of the ripple. This occurs in both closed- and open-bubble situations (any angle

of current incidence)¹⁵. However, at angles of current incidence $<40^\circ$ – 45° an open separation bubble having helical flow should extend beyond the end of the bedform and facilitate elongation of the LTR through convergent bedload transport. As the LTRs propagate, the original flow obstacle may be modified and eventually be removed by erosion. As with the original mound-and-tail, the LTRs parallel mean flow direction.

Of the four factors requisite to this model, the presence of a significant mean flow and superimposed variability are reasonably well documented from bottom photographs¹⁶ and current measurements¹⁴. Current measurements at 5,022 m within the LTR area (Fig. 1) show a mean velocity of 28 cm s^{-1} towards 253° , nearly parallel to the regional contours ($\sim 245^\circ$) over a 15-day period. Currents vary a total of about 110° in direction (190° – 300°); for current speeds in excess of 50 cm s^{-1} , the total variation is about 60° (240° – 300°). Longer-term records at 4,950 m (215 days) and 4,770 m (230 days) have considerably more variation in direction and lower maximum speeds ($<40 \text{ cm s}^{-1}$), but they still show a significant mean velocity of 7 – 8 cm s^{-1} directed towards about 250° (ref. 14).

The degree to which ripple construction results from bedload or suspended load transport is unclear. Bedload transport is suggested by better development of LTRs in the slightly coarser sediment on the south-west part of the Nova Scotian rise, but the texture and the LTR development could be independent responses to the particular current regime in this area. A mechanism whereby suspended sediment is 'plastered' onto the bedform cannot be ruled out. Finally, the requirement for the presence of a flow obstacle to initiate LTR formation is supported by numerous photographs showing biological mounds at the head of LTRs as well as in varying states of elongation by currents (Fig. 2). I cannot absolutely define any of the time scales involved, but in light of the variability recorded in current measurements it is reasonable to suppose that the ripples take at least months to years to form beneath the mean flow and that the short-term variations (and small-scale lineations) occur over periods of days to weeks.

This model contrasts with one proposed by Flood⁷ for LTRs in the Blake–Bahama Basin. He suggested that LTRs may form by very rapid deposition beneath short-lived, high-velocity current events. This explanation is unlikely for the Nova Scotian continental rise because such current events commonly are not oriented parallel to the contours (deviations up to 55° , ref. 14) while the LTRs almost always are parallel. Furthermore, photographs from the Nova Scotian rise show that many LTRs consist of highly consolidated mud, often gouged at oblique angles by tool marks made by large bedload-transported particles. If the LTRs were evanescent bedforms, it is doubtful that they would achieve such strong consolidation. In the Blake–Bahama Basin, numerous biologically produced mounds of scale similar to the LTRs are observed in Flood's photographs, and the bedforms therefore may have an origin similar to those on the Nova Scotian rise.

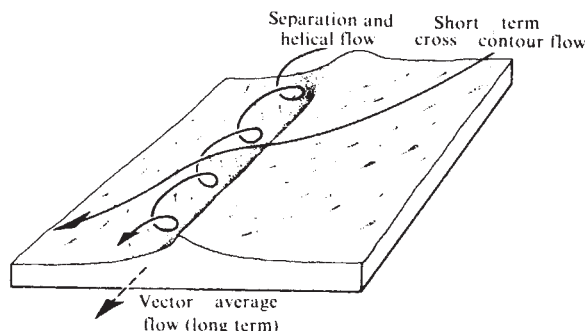


Fig. 3 Sketch showing propagation of LTR in direction of mean flow behind a biological mound. Instantaneous current oblique ($<45^\circ$) to LTR and to mean flow stimulates helical circulation in separation bubble at left side of LTR. Alternations in current direction shift separation bubble between left and right sides of LTR, and they average to the long-term mean flow.

The qualitative model I have proposed does not deal with a host of important subtleties (such as the distribution of bed shear stress) in the nature of interaction between the boundary-layer flow and the bedforms. However, it presents a working hypothesis that agrees with available data and that can be tested by the future HEBBLE experiments and in laboratory conditions. The model predicts that longitudinal triangular ripples are characteristic of a specific set of flow conditions, namely significant short-term variability superimposed on a well-defined and geologically significant mean flow. Such conditions are most likely to be met on low regional slopes where deep boundary currents are established, especially along the western margins of ocean basins.

This work was supported by ONR contract N00014-79-C-0071 at Woods Hole Oceanographic Institution. I thank L. Sullivan and P. Bruchhausen for assistance with the photographic systems, which were supported by ONR Contract N00014-75-C-0210 at Lamont-Doherty Geological Observatory. I. N. McCave, D. G. Aubrey and C. D. Hollister made useful comments. Contribution no. 4984 of Woods Hole Oceanographic Institution.

Received 26 October 1981; accepted 17 February 1982.

1. Hersey, J. B. (ed.) *Deep-Sea Photography* (John Hopkins Press, Baltimore, 1967).
2. Heezen, B. C., Hollister, C. D. & Ruddiman, W. F. *Science* **152**, 502–508 (1966).
3. Tucholke, B. E., Wright, W. R. & Hollister, C. D. *Deep-Sea Res.* **20**, 973–997 (1973).
4. Hollister, C. D., Southard, J. B., Flood, R. D. & Lonsdale, P. F. in *The Benthic Boundary Layer* (ed. McCave, I. N.) 183–204 (Plenum, New York, 1976).
5. Zimmerman, H. B. J. *geophys. Res.* **76**, 5865–5876 (1971).
6. Hollister, C. D., Flood, R. D., Johnson, D. A., Lonsdale, P. & Southard, J. B. *Geology* **2**, 395–400 (1974).
7. Flood, R. D. *Mar. Geol.* **39**, M13–M20 (1981).
8. Ewing, M. & Mouzo, F. *Proc. natn. Acad. Sci. U.S.A.* **61**, 787–793 (1968).
9. Hollister, C. D., McCave, I. N., Laine, E. P., Lonsdale, P. F. & Richardson, M. J. *Deep-Sea Res.* (submitted).
10. Heezen, B. C. & Hollister, C. D. *Mar. Geol.* **1**, 141–74 (1964).
11. Hollister, C. D., Nowell, A. R. M. & Smith, J. D. 3rd A. Rep. *High Energy Benthic Boundary Layer Experiment* (Woods Hole Oceanographic Institution Ref. 80-32, 1980).
12. Thorndike, E. M. *Deep-Sea Res.* **5**, 234–237 (1959).
13. Le Pichon, X., Auzende, J. M., Pautot, G., Monti, S. & Francheteau, J. *Nature* **230**, 110–111 (1971).
14. Richardson, M. J., Wimbush, M. & Mayer, L. A. *Science* **213**, 887–888 (1981).
15. Allen, J. R. L. *Current Ripples, Their Relation to Patterns of Water and Sediment Motion* (North-Holland, Amsterdam, 1968).
16. Tucholke, B. E., Hollister, C. D., Biscaye, P. E. & Gardner, W. D. *EOS* **61**, 1014 (1980).

Drifting buoy trajectories in the North Atlantic Current

W. Krauss & J. Meincke

Institut für Meereskunde, University of Kiel, D 2300 Kiel 1, FRG

Despite intensive research, there are still large gaps in our knowledge of the North Atlantic Current which is considered to be the link between the Gulf Stream off the east coast of North America and the anomalously warm surface waters off northwestern Europe. In the classical scheme given in Fig. 1, which was constructed by Dietrich¹ on the basis of data obtained during 1957–58, the current is seen to comprise several branches. Two important difficulties arise from this interpretation. First, the data for the area east of Newfoundland are not sufficient to define the causes of why the branching of the North Atlantic Current is stationary in its source area over the slope off Grand Banks². Second, rapid surveys of the oceanic temperature structure by means of expendable sondes and, increasingly, of long-term moored instrumentation and satellite-tracked drifting buoys³ have revealed that the oceanic current field is dominated by mesoscale eddies and meanders rather than by large-scale steady flows. We report here two experiments using satellite-tracked buoys which address these difficulties. Our results clarify the pathway of the North Atlantic Current and confirm that the current indeed consists of mesoscale eddies and meanders.

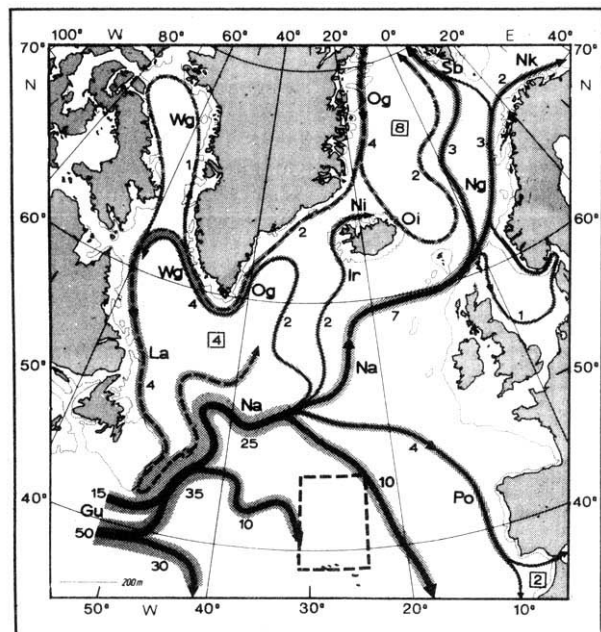


Fig. 1 The branching of the North Atlantic Current (Na) as inferred from the hydrographical data obtained during the International Geophysical Year 1957–58 (see refs 1, 2). The numbers indicate transports in the layer 0–1,000 m, given in $10^6 \text{ m}^3 \text{ s}^{-1}$. The dashed box depicts the area shown in Fig. 4.

In summer 1981, we carried out two experiments. The first was designed to verify the branching of the North Atlantic Current east of Newfoundland. Ten satellite-tracked buoys with drogues at 10 m depth were deployed by the Canadian RV *Baffin* between 5–8 and 14–16 May 1981 along two lines normal to the strong frontal zone south-east of Grand Banks. The lines were about 130 km apart. The buoys were launched at about 50 km intervals along these lines, starting in the zone of major horizontal gradients and running out into warm water. According to Fig. 1, this deployment would be predicted to release the buoys into the waters of the 'central' branch turning towards the Azores, into the 'northern' branch deviating into the southern Labrador Sea before crossing the Mid-Atlantic Ridge near 52°N and into the region of the anticyclonic gyre, located, according to Mann⁴ and Fig. 1, between the central branch and the Gulf Stream termination area.

Figure 2a shows the trajectories of two groups of buoys. Drifters 3509 and 3507 were released at 12.2 and 10.0°C sea-surface temperature in the most northwestern positions of the two launching lines. Both of them followed the northern branch of the North Atlantic Current; Buoy 3507 was trapped by the Flemish Cap for about 3 months before continuing northwards. The second group consisted of drifters 3500, 3501, 3503 and 3506, which were released at the most southeastern ends of the two launching lines at temperatures of 15.5 , 16.5 , 16.0 and 16.5°C respectively. Buoy 3506 failed after a few weeks, while the other three drifters remained stagnant in the gyre area and could not escape until at least mid-September. The trajectories of a third group are shown in Fig. 2b. It includes buoys 3502, 3504, 3505 and 3508, which were deployed between the two groups shown in Fig. 2a at water temperatures of 15.3 , 14.7 , 16.5 and 16.8°C respectively. All four buoys entered the central branch, and on nearing the Gauss Seamount group, two of them passed it to the north and two to the south. They continued to meander towards the Mid-Atlantic Ridge.

It is clear from Fig. 2a, b that the amplitudes of meandering and eddying are large enough for the trajectories of the different groups of drifters to overlap. For example, at around 2 June drifters 3504 and 3508 passed through the 'stagnation' area on their southern paths around the Gauss Seamounts; they had thus crossed the trajectory of drifter 3501 65 days earlier and would cross that of drifter 3500 8 days later. Drifter 3501 cut across the 1957–58 position of the current branch and the related front south-east of the Grand Banks without any significant change in the surface temperature readings. Sixty days later it also cut across the trajectories of the four drifters that finally followed the central branch. The existence of branches and a gyre (or stagnation area), as indicated in Fig. 2a, b, would suggest at least 200 km of spatial variability in the pathway of the current. A preliminary analysis suggests that wind did not significantly affect the trajectories during the observation period. First, the drogue-load sensors that were operational indicated no losses of drogues during the experiment. Second, winds were generally weak and variable during the first 2 months and even when they increased over the area between 30° and 50°W and 40° and 50°N from the south-west for two extended periods in late July and August, they did not cause a uniform displacement of the buoy field.

A few studies of the meandering and eddying of drifting buoys in the western³ and eastern⁵ Atlantic have provided evidence for a relationship between buoy trajectories and individual hydrographical features like Gulf Stream meanders or eddies and rings. We have now investigated this relationship

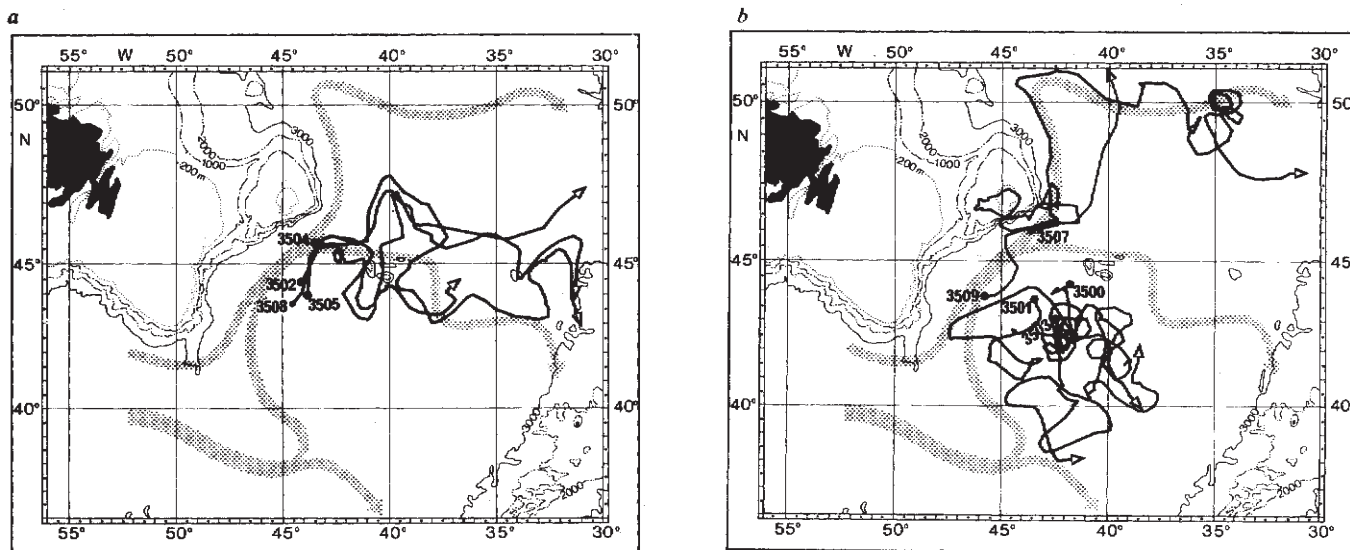


Fig. 2 a, b, Five-month trajectories of drifters drogued at 10 m depth and released mid-May 1981. The start of the trajectories is denoted by a dot and the number of the drifting buoy. The stippled areas correspond to the current branches shown in Fig. 1. Depths exceeding 3,000 m not contoured.

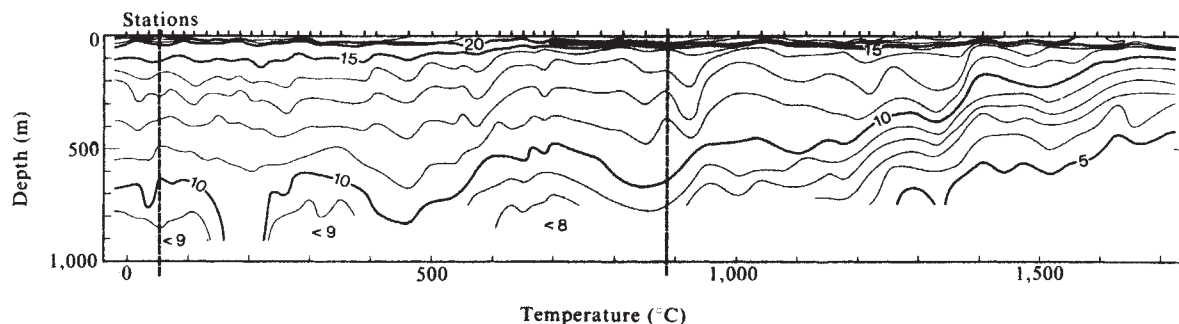


Fig. 3 Temperature section along the eastern flank of the Mid-Atlantic Ridge north of the Azores obtained on 12–16 August 1981. The portion of the section between the dashed lines corresponds to the meridional extent of Fig. 4. The small map shows locations in relation to the Mid-Atlantic Ridge (stippled area).

to the one presented in Fig. 4 covers all this area. The buoys are still drifting across the North Atlantic and are expected to arrive off the European coasts in spring 1982.

Drogued drifters provide a useful tool with which to investigate the kinematics of the current system, especially if simultaneous surveys of the hydrographical structures are possible. The extent to which satellite monitoring of sea-surface temperature pattern can help interpretation of buoy trajectories is not yet clear. IR satellite images of the region shown in Fig. 4 do not reveal any eddy field for the southern and the eastern

for a 700×400 km area over the Mid-Atlantic Ridge north of the Azores in August and September 1981. Figure 1 shows the location of this area between the northern and the central branch of the North Atlantic Current. A three-dimensional survey of the temperature and salinity fields was carried out at the same time as nine drifters drogued at 30 and 100 m depth were followed. The hydrographical setting in Fig. 3 shows the area to be dominated by mesoscale features. It also indicates that the drogues of the drifters were reaching down through the surface mixed layer. Figure 4 displays the depth of the 10°C isotherm, which is representative of the depth of the main thermocline. It is based on 363 XBT and 49 CTD stations with an 'average' spacing of ~ 20 km. Four cold and three warm eddies have been found with an average diameter of 150 km. The trajectories of the buoys superimposed in Fig. 4 indicate a geostrophic flow around the eddies as a first approximation. The close correlation between the trajectories from the uppermost layers and the hydrographical structure at the depth of the main thermocline indicates that the eddy field dominates the flow in the warm water sphere. This result is confirmed by current meter records obtained by a moored vertical array. The current shear observed in the depth range from below the mixed layer to 662 m depth corresponds to that expected from the horizontal temperature gradients shown in Fig. 3.

It took 25 days to complete the hydrographical survey, including a repeat survey of the central and the northeastern region of the area shown in Fig. 4. As the trajectories presented in Fig. 4 are for a period 9 days longer, the eddy field must constitute a quasi-synoptic distribution. This is a basic requirement for most of the further analysis to be performed on this data set.

Another set of drifting buoys was released in August 1981 over the Mid-Atlantic Ridge between 46°N and Gibbs Fracture Zone at 52°N . It seems from the paths that an eddy field similar

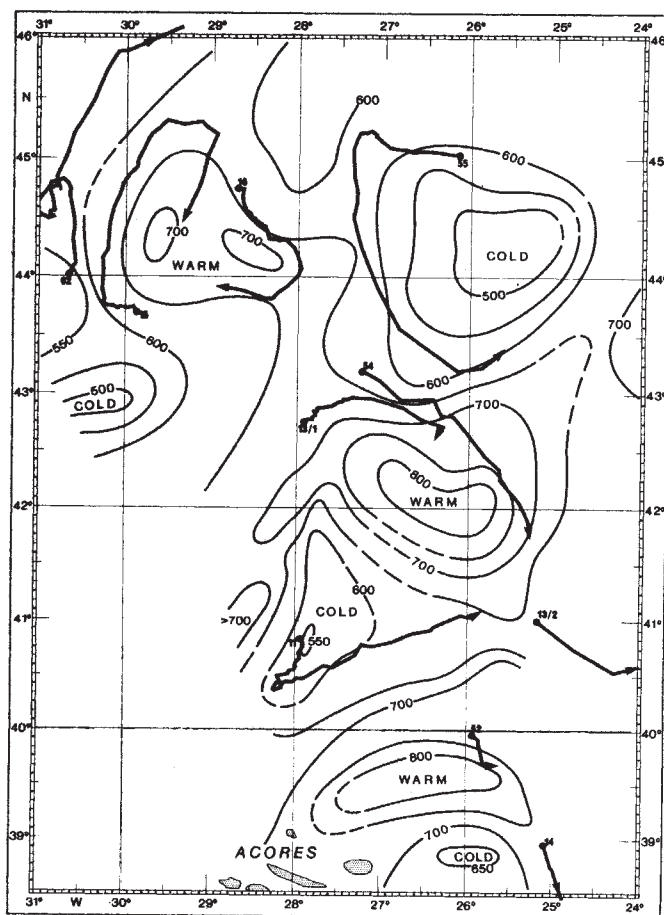


Fig. 4 The topography of the 10°C isotherms and the trajectories of drifters drogued at 30 and 100 m depth. The start of the trajectories is denoted by a dot and the number of the drifter. The hydrographic survey lasted from 12 August to 6 September 1981. The trajectories are shown for the period 12 August to 15 September, with the exception of drifters 14 and 52, which started on 5 September. For general location of the area see Figs 1, 3.

part of the study area, although some features can be identified in the northwestern part.

We acknowledge the active participation of A. Clarke, E. Fahrback, G. Hardtke, J. Stahlmann and A. Sy in this work.

Received 14 December 1981, accepted 1 March 1982.

1. Dietrich, G., Kalle, K., Krauss, W. & Siedler, G. *Allgemeine Meereskunde* 3rd edn, 593 (Gebr. Borntraeger, Berlin, 1975).
2. Mann, C. R. *Proc. R. Soc. Edinb.* B72, 341–349 (1972).
3. Richardson, P. L. *J. Phys. Oceanogr.* 10, 90–104 (1980).
4. Mann, C. R. *Deep-Sea Res.* 14, 337–359 (1967).
5. Madelain, F. & Kerut, E. G. *Oceanol. Acta* 1, 159–168 (1978).

Cooperation and competition within coalitions of male lions: kin selection or game theory?

Craig Packer & Anne E. Pusey

Allee Laboratory of Animal Behavior,
5712 South Ingleside, Chicago, Illinois 60637, USA

Male lions form cooperative coalitions which compete against other coalitions for exclusive access to female groups^{1,2}. This cooperation and the apparently low level of intra-coalition competition over oestrous females, have been considered to be due to the close genetic relatedness of the males in the coalition^{1–4}. However, we now present evidence that breeding coalitions of male lions include non-relatives much more commonly than was generally supposed, that intra-coalition competition over females is widespread and that kinship is not the primary factor determining levels of competition.

Between July 1978 and May 1981 we studied the population of lions in the Serengeti, Tanzania, described by Schaller⁵, and all the lions resident on the floor of Ngorongoro Crater. Lion prides are stable social groups composed of 2–18 adult females, their dependent offspring and a coalition of 1–7 adult males (which can simultaneously control more than one group of females). Genealogical records have been maintained since 1966 for two prides in the Serengeti⁶ and since 1974 for another 13 prides in the Serengeti and Ngorongoro². We made daily censuses of as many prides as possible and recorded the incidence of fresh wounds, the consort partners and mating activity of each male, and the presence of non-consorting males ('rivals') within 200 m of each consort pair. We collected over 500 h of behavioural data on consort pairs during 2 h watches at dawn and dusk.

Theoretical analyses of male coalitions have assumed that they are always composed of relatives^{1,2}. However, there are examples of unrelated males becoming companions^{2,5}, and our data show that 42% of breeding coalitions of known origins

contained non-relatives (Table 1). This proportion is higher than the 10% reported by Bygott *et al.*², because their data included cohorts that had not yet gained a pride (J. D. Bygott, personal communication). There is considerable mortality among subadult and nomadic males so that by the time they gain a pride their initial companions may have been lost and new ones found.

Competition for individual oestrous females between males from the same coalition consists primarily of competition for temporary 'ownership' of the female. When a female in a pride comes into oestrus, the first male that encounters her forms a consortship with her and this 'confers temporary dominance on the consorting male'^{5,6}. The consorting male maintains proximity (usually <1 m) to the female, herds her and prevents other males of his coalition from moving too close to her. Behavioural oestrus (days with mating) lasts about 4 days, during which time copulation occurs once every 25 min; the inter-oestrus interval is about 16 days⁷. The male sometimes guards the female for 1 or 2 days before mating and for up to 6 days after mating has ceased⁷. Males occasionally 'mistake' a potentially oestrous female, guarding her for a day or two and then leaving her after she fails to come into oestrus.

As both coalitions containing relatives and those containing non-relatives were common, it was possible to determine whether kinship affected intra-coalition aggression over oestrous females. Threats without contact by consorting males towards other males were common and occurred mainly when the other males were within 15 m of the consorting male's female (84%, $n = 63$ threats). Most of these threats (53%) were the direct result of the female attempting to move closer to rivals (or other consorting males). Kinship had no effect on the frequency of threats by consorting males to rivals (calculated per minute spent by a rival within 15 m of a consorted female and based on six consorting males whose unrelated companion(s) spent any time within 15 m of his consort partner, and on 12 males whose related companion(s) spent time within 15 m; $U = 36$, $P > 0.50$), or the time spent by rivals within 15 m of the consorted female (based on 10 males who had unrelated consorting companions, and 19 males who had related consorting companions; $U = 76$, $P > 0.20$). The fact that rivals spent any time as close as 15 m to the consort pair was due primarily to the behaviour of the female. Their presence at more moderate distances, however, depended on the behaviour of the rival, and rivals were more likely to be within 200 m of a consorted oestrous female as the availability of other possibly oestrous females declined (Fig. 1).

More serious intra-group fights over females included slapping and biting and frequently resulted in wounding of one or both males. Such fights were observed nine times and could be inferred another five times from fresh wounds on one or both males together with a change of consort partners. Serious fights were no more common between non-relatives than between relatives ($P > 0.50$); rather, they were context specific: of 13 fights where the context was known or could be inferred, eight occurred when 'ownership' of the female was undecided or unclear and four involved two consorting males. Ownership was 'undecided' when two males simultaneously came into the vicinity of an unconsorted, potentially oestrous female; and was 'unclear' when the consorting male moved further from the female than was the rival. There was usually a race between males to arrive first at a female. On arrival, the loser would defer to the winner though in subsequent consortships the previous consorting male might be a rival to the new consorting male (also see ref. 5). Lion males can consort simultaneously with two females and females being consorted by different males occasionally tried to come into close proximity to each other. In such cases one consorting male sometimes tried to take over the other's female.

Males also competed in a less direct manner to be the owner of a female when she came into oestrus. Males guarded non-oestrous females, probably in anticipation of the female's oestrus, and this was most common when the availability of

Table 1 The composition and origins of the 20 male coalitions having tenure in prides in the Serengeti and Ngorongoro during 1978–81

Coalition sizes	Origins
6, 5, 5, 4, 3, 2, 2	Unknown
3	Pair of unknown origins with son
7, 4, 4, 4, 3	Cohort from same natal pride
3, 2	Full siblings
3, 2, 2, 2	All unrelated
3	Sibling pair with unrelated third

Males were considered relatives if they came from the same natal pride or joined their probable fathers, and non-relatives if they were known not to have come from the same natal pride or if they had associated with a variety of male partners for over a year before coming together (consistent with the behaviour of males known to be from different prides and completely unlike that of males of the same natal pride).

oestrous females was lowest. When only one female in a pride was being consorted, the consorting male was more likely merely to be guarding a non-oestrous female than on days when several females were being consorted ($T = 40$, $n = 19$ males observed consorting at least twice in each condition, $P < 0.05$).

Males in Ngorongoro Crater guarded non-oestrous females more frequently than did males in the Serengeti. When only one female was consorted, the female was not in oestrus on 80% of occasions in the Crater (based on the behaviour of 11 consorting males) but on only 42% of occasions in the Serengeti ($n = 24$, $U = 69$, $P < 0.025$). Guarding of non-oestrous females may be more common in the Crater because it is less costly. Consort pairs rarely hunt and therefore feed infrequently⁵, but prey density in the Crater is so high and pride ranges so small⁶ that consort pairs can more easily join pride mates at kills.

Because fights over oestrous females are so rarely observed, males of a coalition have been said to 'share' females and to do so because they are relatives (see refs 1, 2). However, two factors should be taken into account when attempting to measure differential reproductive success among males of a coalition. First, females in a pride tend to come into oestrus at the same time^{5,6}. In our study, on the first day of 43% of oestrous periods other females in the pride were also in oestrus ($n = 150$). Second, females tend to move to additional mating partners after their first mating partner loses interest in them at the end of oestrus⁷. Subsequent partners show only a brief interest in the female and females seek additional partners most often when their fertility is lowest⁷. Thus calculating the relative mating activity of males without controlling for either the number of females in oestrus or the order in which males consorted with a particular female (as in refs 1, 2) biases against finding any differences in reproductive success.

The extent of differential male reproductive success was measured in 11 coalitions using data that eliminated these biases (Table 2). Coalitions consisting solely of relatives did not 'share' females more equally than did coalitions containing non-relatives. However, differences in competitive ability did result in differential male reproductive success. Disparities in consorting success were greater in coalitions in which the males were of markedly different ages or sizes than in those where all the members were evenly matched (Table 2). In all six of the former coalitions, the non-prime (either very young or old) and small individuals consorted less frequently with oestrous females than did their larger or more vigorous companions.

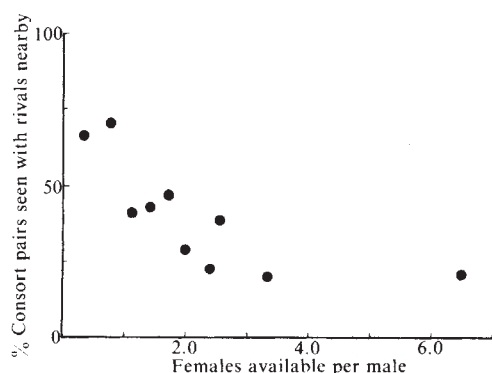


Fig. 1 The proportion of sightings ($n = 222$) in which non-consorting males ('rivals') were seen within 200 m of consort pairs plotted as a function of the number of potentially oestrous females available to those males elsewhere (number of unconsorted females not pregnant or with cubs/number of non-consorting males). Only data taken from coalitions which had sole control of a pride are used and points were taken at least 4 days apart. Eight of nine coalitions showed trends similar to the overall result ($P < 0.04$) and there was no difference between the Serengeti and Ngorongoro, so all data were pooled and cells combined to give a $k \times r$ matrix with sufficient sample size to test for heterogeneity ($\chi^2 = 23.87$, 9 d.f., $P < 0.01$). The trend for rivals to be present more often when fewer females are available elsewhere is significant ($r_s = 0.89$, $n = 10$, $P < 0.01$).

Table 2 Extent of differential mating success in each coalition

Mating success of respective males	Includes non-relatives?	All same age and size?	I_T
3, 0, 0	Unknown	No	2.00
3, 0, 0	Unknown	No	2.00
5, 5, 0, 0, 0	Unknown	No	1.50
4, 0	Yes	No	1.00
2, 0	Yes	No	1.00
3, 3, 0, 0	Unknown	No	1.00
7, 4, 3, 2, 2, 1, 1	No	Yes	0.47
7, 4, 2, 2	No	Yes	0.30
3, 2, 1	No	Yes	0.17
2, 1, 1	Yes	Yes	0.13
2, 2	Yes	Yes	0

Each line gives data on a different male coalition. For each male in each coalition, the number of oestrous periods is given in which he consorted with a female who was the only oestrous female available to that coalition on the day when she was first observed mating. Only the first male to mate with the female in each of these periods is considered, unless a male was seen to win the female from another male in a fight—in these cases the victor was also included. Only coalitions in which the average number of such periods is one or more per male are included. For each coalition a measure of differential mating success, I_T (ref. 14), was calculated. Large values of I_T indicate unequal mating success; smaller values indicate evenly distributed mating success. Coalitions containing members of different size or age showed significantly more unequal mating success than coalitions with equally aged and sized members ($n_1 = 6$, $n_2 = 5$, $U = 0$, $P < 0.01$), but there was no significant difference between coalitions consisting solely of relatives or including non-relatives ($n_1 = 3$, $n_2 = 4$, $U = 6$, $P > 0.50$).

The high proportion of coalitions containing non-relatives is not surprising. Because singleton males almost never gain control of a pride and coalitions of three or more father more surviving offspring per male than do smaller coalitions², selection will favour singletons and pairs forming coalitions with additional males. As females are frequently in oestrus simultaneously, even subordinate males often have access to females.

Males would benefit even by cooperating with non-relatives but should prefer relatives as partners^{9,10}. We know of no case where a male with related companions left them for an unrelated companion, although there were cases where additional relatives were not accepted. A group of three relatives prevented three younger relatives from joining them, and incorporation of sons into the coalition was known or suspected only when the number of older males was one or two ($n = 4$ coalitions). In coalitions of three or more, incorporation of sons has never been seen, even though such large coalitions are more likely to maintain tenure long enough to father surviving sons². Because large coalitions can control prides without additional companions, they would be expected to exclude additional members when an added male would reduce each male's reproductive success (but not when the exclusion results in a loss in inclusive fitness). This is especially important if the new male would become disproportionately successful as in the case of ageing males being joined by males nearing their prime.

Costs to males of direct competition for oestrous females can be high: one-to-one fights typical of such encounters often result in wounds to the face and eyes and sometimes in blinding⁵. Even in a gang attack on a single individual, the lone animal can wound several of its opponents. Furthermore, the loss of a companion through fighting may shorten tenure in the pride¹. Game theoretical analysis predicts that when costs of fighting are high, contests may be settled 'conventionally' through recognition of asymmetries such as 'owner versus rival' or 'large versus small', rather than through overt aggression¹¹⁻¹³. As differences in size or vigour exist in only about one-third of male coalitions (7 of 20), the 'respect' of ownership in lions is particularly important. Males were seen to compete for females indirectly by anticipating oestrus in a female in order to be the 'owner' when she eventually came into oestrus and they did

this more often when costs of guarding and availability of oestrous females were lowest. Fights were virtually restricted to occasions when ownership was unclear or when two consorting males were in close proximity.

In conclusion, the low levels of aggression observed within coalitions of male lions, which had been ascribed to kin selection, are not affected by the degree of genetic relatedness of the males and may be better understood in terms of game theory.

We thank the H. F. Guggenheim Foundation, the Royal Society of Great Britain (to A.E.P.) and the National Geographic Society for support of fieldwork; the Government of Tanzania and the Serengeti Research Institute for permission and facilities; B. C. R. Bertram, J. D. Bygott and J. P. Hanby for data and for ensuring the continuity of the lion study; and J. Altmann, S. A. Altmann, S. J. Arnold, P. H. Harvey and J. Silk for comments. C.P. is supported by NIMH training grant MH15181.

Received 11 November 1981; accepted 12 February 1982.

1. Bertram, B. C. R. in *Growing Points in Ethology* (eds Bateson, P. P. G. & Hinde, R. A.) (Cambridge University Press, 1976).
2. Bygott, J. D., Bertram, B. C. R. & Hanby, J. P. *Nature* **282**, 839–841 (1979).
3. Alcock, J. *Animal Behavior: An Evolutionary Approach* 2nd edn (Sinauer, Sunderland, Massachusetts, 1979).
4. Wittenberger, J. F. *Animal Social Behavior* (Duxbury, Boston, 1981).
5. Schaller, G. B. *The Serengeti Lion* (Chicago University Press, 1972).
6. Bertram, B. C. R. *J. Zool., Lond.* **177**, 463–482 (1975).
7. Packer, C. & Pusey, A. E. *Am. Nat.* (submitted).
8. Elliott, J. P. & Cowan, I. M. *Can. J. Zool.* **56**, 1726–1734 (1978).
9. Hamilton, W. D. *J. theor. Biol.* **50**, 1–52 (1964).
10. West Eberhard, M. J. *Q. Rev. Biol.* **50**, 1–34 (1975).
11. Maynard Smith, J. & Parker, G. A. *Anim. Behav.* **24**, 159–175 (1976).
12. Maynard Smith, J. *Proc. R. Soc. B205*, 475–488 (1979).
13. Hammerstein, P. *Anim. Behav.* **29**, 193–205 (1981).
14. Arnold, S. J. & Wade, M. J. *Evolution* (in the press).

Properties of the Ca^{2+} -activated K^+ channel in one-step inside-out vesicles from human red cell membranes

Virgilio L. Lew, Shmuel Muallem*
& Carol A. Seymour

Physiological Laboratory, University of Cambridge,
Cambridge CB2 3EG, UK

The presence of a calcium-activated potassium channel in some mammalian red cell membranes makes them useful experimental models in which to study the properties of similar mechanisms believed to be involved in the control of membrane potential and conductance, at rest and during activity, in many other cells. However, vesicles prepared from human red cell membranes by the method of Steck *et al.*¹, whether inside-out (IOVs) or right-side out (ROVs), either failed to show any Ca^{2+} -activated component of $\text{K}^+(\text{Rb}^+)$ fluxes² or showed only a very reduced calcium sensitivity and K^+/Na^+ selectivity³. This is surprising because there is good functional preservation of other transport mechanisms, such as the anion carrier⁴, Ca^{2+} pump⁵ and Na^+ pump⁶. Our failure to confirm reports⁷ that the calcium response could be restored in 'silent' IOVs by addition of protein concentrates from red cell lysates prompted the search for and discovery of vesiculation procedures which produced ion-tight IOVs in a single step, with minimum loss of membrane components and transport properties^{8–11}. We report here (1) the conditions which favour preservation or loss of the Ca^{2+} -activated component of the ^{86}Rb efflux from one-step IOVs, (2) an approximate estimate of the number of Ca^{2+} -activated K^+ channels per red cell, and (3) that individual Ca^{2+} -activated K^+ channels respond in an all or nothing fashion to Ca^{2+} activation and differ in their threshold sensitivity to ionized calcium.

The IOVs used in these experiments were prepared by the one-step method and loaded with $^{86}\text{Rb}^+$ as detailed in Fig. 1 legend. Preliminary experiments had indicated that when the one-step IOVs are resuspended in media which can sustain the activity of the Na^+ pump (see Fig. 1 legend), the $^{86}\text{Rb}^+$ release caused by ATP addition exposes the total $^{86}\text{Rb}^+$ space sealed within IOVs that contain Na^+ pumps. Figures 1 and 2 compare this Na^+ -pump IOV space with that accessible to Ca^{2+} -activated channels, in conditions which optimally preserve (Fig. 1) or partially inhibit (Fig. 2) the Ca^{2+} -dependent $^{86}\text{Rb}^+$ release.

In 11 experiments similar to that of Fig. 1, using red cells from three different donors, the ratio, r , of the $^{86}\text{Rb}^+$ space accessible to Ca^{2+} -activated channels to the $^{86}\text{Rb}^+$ space accessible to Na^+ pumps varied between 0.86 and 1.02. With an average of 500–1,500 vesicles of non-uniform size per cell¹¹ and 100–200 Na^+ pumps per cell^{12,13}, the probability of a vesicle having more than one pump will be rather low. However, most of the $^{86}\text{Rb}^+$ released will be from the larger vesicles which contain a larger fraction of the $^{86}\text{Rb}^+$ -sealed vesicular volume¹¹. These vesicles are more likely to have both pumps and channels, thus causing more overlap than expected had all the vesicles been of uniform size. If the surface density of pumps and channels is assumed to be uniform, the similarity of $^{86}\text{Rb}^+$ space accessible to Na^+ pumps and K^+ channels could reflect identity of the two mechanisms¹⁴, structural linkage or simply numerical parity. The first two possibilities are unlikely because of the additional $^{86}\text{Rb}^+$ -releasing effects of Ca^{2+} and ATP shown in Figs 1b and 3, which suggest that there are vesicles with either one or other of the two transporting systems. If the assumptions made here were correct, the number of Ca^{2+} -sensitive K^+ channels, like that of Na^+ pumps^{12,13}, would be in the range of 100–200 per cell.

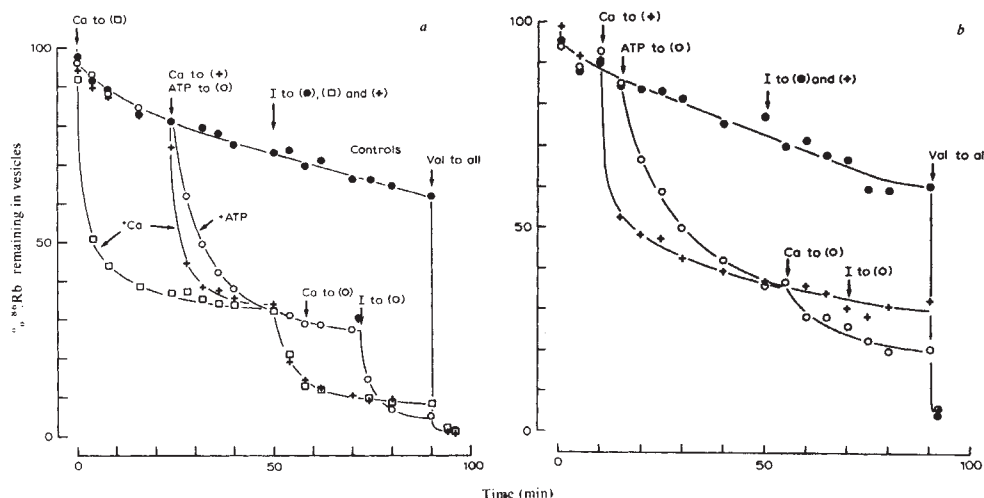
The experiment of Fig. 2 represents the conditions which cause relative channel inhibition. Lysis and vesiculation at high pH and in the presence of EDTA cause a variable reduction in r (0.06–0.65 range, in seven experiments similar to that of Fig. 2). As with the uninhibited channels (Fig. 1), the Ca^{2+} -induced $^{86}\text{Rb}^+$ release also saturated at about 50 μM Ca^{2+} . Addition of concentrated protein extracts from red cell lysates or purified calmodulin¹⁵, as with Steck-type vesicles, failed to restore r and only had minute random effects. Partial inhibition, therefore, seems to represent irreversible channel inactivation. Low pH (7.5) *per se* had no effect on r . The critical condition for optimal channel inactivation seems to be enhanced divalent cation chelation by high pH (8.5) during vesiculation, as if the most vulnerable configuration were only transiently present at this stage. Higher EGTA or EDTA concentrations (0.5–3 mM) interfered with vesiculation itself and could not be tested. Used before or after, but not during, vesiculation, they caused only small and rather variable reductions in r even at pH 7.5, and, as before, only in the absence of added Ca^{2+} or Mg^{2+} . These inhibitory effects of EGTA or EDTA at high pH could perhaps explain the failure to observe Ca^{2+} -activated components of the $\text{K}^+(\text{Rb}^+)$ fluxes in some of the Steck-type vesicle preparations².

The protective effect of Ca^{2+} and Mg^{2+} suggests an interesting alternative interpretation of the Ca^{2+} -sensitivity curves reported in resealed ghosts containing buffered calcium^{16–18}. It may be that part of the reported Ca^{2+} affinity for activation of the channel represents Ca^{2+} affinity for protection against irreversible inhibition by EGTA. This interpretation is supported by preliminary experiments with resealed ghosts containing 1 mM EGTA and $^{86}\text{Rb}^+$. The Ca^{2+} -dependent $^{86}\text{Rb}^+$ release was inhibited by up to 80% when Ca^{2+} was incorporated using the ionophore A23187 only during a final incubation relative to ghosts sealed from the beginning with similar buffered Ca^{2+} concentrations.

In 'one-step' IOVs, with partial or total preservation of functional Ca^{2+} -sensitive K^+ channels not able to be inhibited further, the true Ca^{2+} sensitivity of the K^+ channels may be studied directly. As shown in Fig. 3, the striking finding was that with increasing Ca^{2+} concentrations, progressively more

* Permanent address: Department of Biochemistry, The Weizmann Institute, Rehovot, Israel.

Fig. 1 Effects of ATP and Ca^{2+} on the $^{86}\text{Rb}^+$ release from one-step human red cell IOVs. Red cells from freshly drawn blood were washed four times in about 8 vols of isotonic media containing NaCl + KCl, 150 mM, and Tris-Cl (pH 7.8 at room temperature), 10 mM, packed to about 80% haematocrit and lysed in 40 vols of an ice-cold medium containing Na-HEPES (pH 7.5), 2.5 mM, and Na-EGTA, 0.1 mM. After about 10 min, to allow for completion of lysis, 40-ml aliquots of this suspension were centrifuged for 10 min at 30,000g in the cold (4°C), the ghost pellets resuspended in 2–4 ml of the same lysing medium and incubated at 37°C for 60 min. Spontaneous inside-out vesiculation occurs during the first 15–20 min of this incubation, as described before⁹, but membrane elasticity increases with time¹¹, allowing subsequent fragmentation and resealing of a fraction of the vesicles by vigorous passage through 27-gauge needles. Earlier measurements of $(\text{Ca}^{2+} + \text{Mg}^{2+})\text{ATPase}$ ¹¹ indicated that at least 25% of the original red cell membrane area surrounds IOV-sealed vesicular space. Traces of $^{86}\text{Rb}^+$ (10–50 $\mu\text{Ci ml}^{-1}$) were added before (a) or after (b) vesiculation. Addition before vesiculation secures maximum $^{86}\text{Rb}^+$ trapping within the resealed vesicular space; after vesiculation, incorporation of $^{86}\text{Rb}^+$ was by needling. The $^{86}\text{Rb}^+$ -loaded vesicles (from 1 ml original 'packed' cells) were washed once (to remove extravesicular tracer), and resuspended in 8–12 ml of a medium containing (mM): NaCl, 50; KCl, 10; Na-HEPES (pH 7.5), 25; MgCl_2 , 1.5; EGTA, 0.02 (medium 'A'). The $^{86}\text{Rb}^+$ remaining in the vesicles was measured, at the indicated times after the start of the final incubation (37°C), in 0.1-ml aliquots of this suspension, run through Tris-neutralized Dowex 50-X-8 (200-mesh) columns contained within glass-wool-plugged Pasteur pipettes and eluted with 1.5 ml of ice-cold sucrose (200 mM) directly into counting vials. The results are expressed as per cent $^{86}\text{Rb}^+$ remaining in vesicles relative to the samples at $t=0$. The indicated additions were taken from concentrated stock solutions (concentration in parentheses) to give the following final concentrations (mM): Na-ATP (100 mM), 1.5; A23187 (I) (2 mM in absolute ethanol) 0.01; CaCl_2 (30 mM), 0.07; valinomycin (Val) (1 mM in ethanol) 0.001. The final Ca^{2+} concentration of 0.05 mM used here elicited maximal $^{86}\text{Rb}^+$ release in all conditions (not shown). The average vesicle diameter was $\sim 0.2 \mu\text{m}$. Addition of ionophore A23187 in the presence of Ca^{2+} revealed a substantial $^{86}\text{Rb}^+$ space contained within ROVs (a). This compartment did not incorporate much tracer by needling (b). Establishing the proportions of ROVs to IOVs by assaying acetylcholinesterase activity as in multistep vesicles, does not give reliable values in one-step IOVs because vesiculation is performed without washing and the enzyme may easily reassociate with both sides of the membrane. Ca^{2+} addition after ATP-induced $^{86}\text{Rb}^+$ depletion had no measurable effect in a but caused further $^{86}\text{Rb}^+$ loss in b. In a, where all the resealed vesicular space was labelled, a disproportionate fraction of this space may be contained within the larger vesicles. Their chance of having more than one pump or channel may be higher than in the smaller labelled vesicles of b and this may explain the larger overlap of $^{86}\text{Rb}^+$ space available to both mechanisms without necessarily postulating identity or link (see text).



vesicles released their $^{86}\text{Rb}^+$ at the same maximum relative rate, instead of, as expected, all the channel-containing vesicles releasing their $^{86}\text{Rb}^+$ at increasingly higher rates. The ratio between the rate measured over the first 5 min and the steady-state difference remained virtually constant at all Ca^{2+} concentrations (see Fig. 3 inset).

This means that Ca^{2+} activation is all or nothing for each channel and that the channels differ in their threshold Ca^{2+} sensitivity. The insert in Fig. 3, therefore, does not report any true or apparent Ca^{2+} affinity, but the distribution of Ca^{2+} thresholds in the channel population of this vesicle preparation. In three other similar experiments, there was some variation in the threshold distribution curve, with channels still being activated in the 2–10 μM Ca^{2+} range, but always in an all or

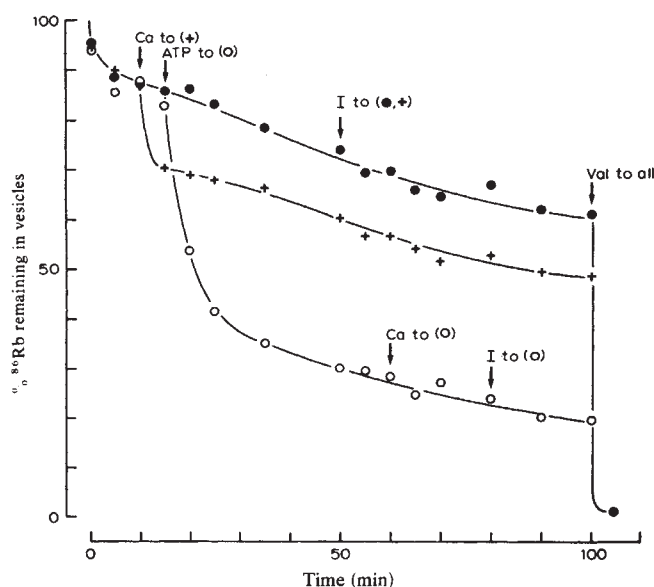


Fig. 2 Inhibition of Ca^{2+} -induced relative to ATP-induced $^{86}\text{Rb}^+$ release in one-step human red cell IOVs. This experiment was identical to that in Fig. 1a except that lysis and vesiculation were performed in a medium containing (mM): Na-HEPES (pH 8.5), 2.5, and Na-EDTA, 0.1. Possible inhibition of the Ca^{2+} -sensitive channels in the ROVs could explain the failure of A23187 to activate further $^{86}\text{Rb}^+$ release as in Fig. 1a. In any case, the contaminant ROV space in this experiment cannot exceed 22% of the total $^{86}\text{Rb}^+$ -resealed vesicular space because at least 78% of it must be contained within IOVs with Na⁺ pumps.

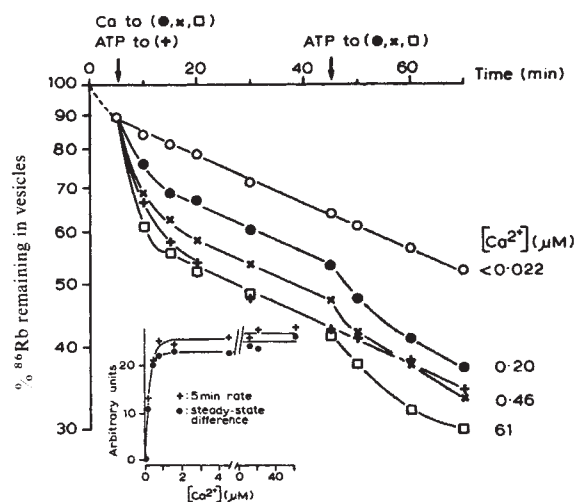


Fig. 3 Ca^{2+} -sensitive $^{86}\text{Rb}^+$ loss from one-step IOVs as a function of the Ca^{2+} concentration. IOVs were prepared as in the experiment of Fig. 1a. The final suspending medium ('A') had a contaminant Ca^{2+} concentration (as determined with Ca^{2+} electrodes) of 7 and 100 μM of added EGTA. CaCl_2 was added to give the calculated concentrations of Ca^{2+} shown. ATP addition as in the experiment of Fig. 1. In three other experiments using medium A and 1 mM EGTA, the threshold distribution curve (see inset) was shifted slightly to the right (app $K_d \approx 2 \mu\text{M}$). Ca^{2+} uptake by the IOVs before ATP addition was less than 2% of the total Ca^{2+} even at the lowest Ca^{2+} concentration (not shown). This rules out Ca^{2+} depletion through vesicle binding as the possible cause of the all or nothing response.

nothing fashion. Apparent changes in Ca^{2+} sensitivity therefore result from shifts in the threshold distribution curve.

If the all or nothing response of each channel reflects the behaviour of the channels in the original cells, a graded increase in $\text{K}^+(\text{Rb}^+)$ flux in intact cells or resealed ghosts, in conditions where the Ca^{2+} concentration is uniformly raised in all the cells, must reflect either activation of only a fraction of the channels in each cell or, if all the channels belonging to the same cell had similar Ca^{2+} thresholds, activation of all the channels in only a fraction of the cells. It is unlikely that cells with few or no channels contribute to the graded response because, when Ca^{2+} is incorporated at high ionophore (A23187) concentrations in intact cells, the entire cell K^+ pool participates in the $^{42}\text{K}^+$ -tracer equilibration process¹⁹. Tracer $\text{K}^+(\text{Rb}^+)$ pools retained within cells with inactive channels have been observed

in intact cells and resealed ghosts²⁰⁻²⁴, although seldom in conditions of known uniformity of Ca^{2+} or Pb^{2+} (ref. 20) content. Such all or none cell responses^{20,24}, however, do not imply all or nothing single channel behaviour as found here, because they can also result from variations in the Ca^{2+} sensitivity of channels belonging to different cells, each channel responding to Ca^{2+} in a graded manner.

Too little precise information exists at present to define the range of possible kinetic models which could explain the present observations. Sharp all or nothing responses at a molecular level may require interacting configurational changes, occluded states or $\text{dCa}^{2+}/\text{dt}$ -sensitive gating mechanisms.

We thank J.D. Cavieres, R. W. Tsien, R. M. Bookchin and I.M. Glynn for helpful discussions, Mrs J. Gray for technical assistance and the MRC and EMBO for funds.

Received 24 September 1981; accepted 25 February 1982.

1. Steck, T. L., Weinstein, R. S., Straus, J. H. & Wallach, D. F. H. *Science* **168**, 255-257 (1970).
2. Grinstein, S. & Rothstein, A. *Biochim. biophys. Acta* **508**, 236-245 (1978).
3. Sze, H. & Solomon, A. K. *Biochim. biophys. Acta* **554**, 180-194 (1979).
4. Rothstein, A., Ramjeesingh, M. & Grinstein, S. *Alfred Benzon Symp.* **14**, 329-340 (1980).
5. Macintyre, J. D. & Green, J. W. *Biochim. biophys. Acta* **510**, 373-377 (1978).
6. Lee, K. H. & Blostein, R. *Nature* **285**, 338-339 (1980).
7. Sarkadi, B., Szebeni, J. & Gardos, G. *Alfred Benzon Symp.* **14**, 220-231 (1980).
8. Lew, V. L., Muallem, S. & Seymour, C. A. *J. Physiol., Lond.* **307**, 36-37P (1980).
9. Lew, V. L. & Seymour, C. A. *J. Physiol., Lond.* **308**, 8-9P (1980).
10. Bookchin, R. M., Raventos, C. & Lew, V. L. *Proc. 5th Int. Conf. Red Cell Metabolism and Function* (ed. Brewer, G. J.) (Liss, New York, in the press).
11. Lew, V. L. & Seymour, C. A. in *Techniques in Lipid & Membrane Biochemistry* (eds Kornberg, H. L. et al.) (Elsevier, Amsterdam, in the press).
12. Dunham, P. B. & Hoffman, J. F. *Proc. natn. Acad. Sci. U.S.A.* **66**, 936-943 (1970).

13. Erdmann, E. & Hasse, W. *J. Physiol., Lond.* **251**, 671-682 (1975).
14. Blum, R. M. & Hoffmann, J. F. *J. Membrane Biol.* **6**, 315-328 (1971).
15. Muallem, S. & Karlish, S. J. D. *FEBS Lett.* **107**, 209-212 (1979).
16. Porzig, H. *J. Physiol., Lond.* **249**, 27-50 (1975).
17. Simons, T. J. B. *J. Physiol., Lond.* **256**, 227-244 (1976).
18. Garcia-Sancho, J., Sanchez, A. & Herreros, B. *Biochim. biophys. Acta* **556**, 118-130 (1979).
19. Lew, V. L. & Ferreira, H. G. *Nature* **263**, 336-338 (1976).
20. Riordan, J. R. & Passow, H. *Biochim. biophys. Acta* **249**, 601-605 (1971).
21. Riordan, J. F. & Passow, H. in *Comparative Physiology* (eds Bolis, L., Schmidt-Nielsen, K. & Maddrell, S. H. P.) 543-581 (North-Holland, Amsterdam, 1973).
22. Lew, V. L. in *Comparative Biochemistry and Physiology of Transport* (eds Bolis, L., Bloch, K., Luria, S. E. & Lynen, F.) 310-316 (North-Holland, Amsterdam, 1974).
23. Knauf, P. A., Riordan, J. F., Schuhmann, B. & Passow, H. in *Comparative Biochemistry and Physiology of Transport* (eds Bolis, L., Bloch, K., Luria, S. E. & Lynen, F.) 305-309 (North-Holland, Amsterdam, 1974).
24. Heinz, A. & Passow, H. *J. Membrane Biol.* **54**, 119-131 (1980).

All-or-none response of the Ca^{2+} -dependent K^+ channel in inside-out vesicles

Javier García-Sancho, Ana Sanchez & Benito Herreros

Departamento de Fisiología y Bioquímica, Facultad de Medicina, Valladolid-5, Spain

In several cell lines, an increase in the intracellular Ca^{2+} concentration elicits an increase in the membrane permeability to $\text{K}^+(\text{Rb}^+)$ by activation of the so-called Ca^{2+} -dependent K^+ channel¹. The human red cell has been widely used to study this transport system². As Ca^{2+} (refs 3, 4) and other possible modulators^{5,6} act on the intracellular side of the membrane, a preparation of everted vesicles would be most useful for investigating such actions. Although several previous studies in inside-out vesicles⁷ have been unsatisfactory⁸⁻¹⁰, a new vesiculation procedure that preserves Ca^{2+} -dependent K^+ transport has recently been reported^{11,12}. Using this preparation we now find that each vesicle responds to Ca^{2+} in an all-or-none fashion, the number of activated (permeable to ^{86}Rb) vesicles increasing with Ca^{2+} concentration between 10^{-8} and 10^{-6} M. This behaviour probably reflects the all-or-none response of individual K^+ channel units. The sensitivity to Ca^{2+} is decreased by Mg^{2+} .

^{86}Rb fluxes were measured at room temperature and in equilibrium exchange conditions. In vesicles loaded with ^{86}Rb during the vesiculation step, the addition of Ca^{2+} to the medium elicited a rapid release of ^{86}Rb , the efflux rate constant returning close to control values (without Ca^{2+}) after 10-20 min. The fraction of vesicular space emptied of ^{86}Rb during this rapid phase increased in a graded stepwise fashion with increasing Ca^{2+} concentrations, being maximal at 10^{-4} M (Fig. 1a). A similar behaviour was observed for the uptake of ^{86}Rb (Fig. 1b). Both the transient nature of these rapid ^{86}Rb fluxes and the graded response suggest that increasing Ca^{2+} concentrations activate an increasing number of vesicles, each vesicle responding in an all or nothing fashion with a fast and full equilibration

of ^{86}Rb . Time-dependent inactivation of the K^+ channel was excluded because vesicles loaded with ^{86}Rb by a 15-min incubation in the presence of submaximal concentrations of Ca^{2+} released most of their radioactivity when transferred to tracer-free medium containing the same Ca^{2+} concentration. On the other hand, if the vesicles were loaded at a Ca^{2+} concentration giving maximal response (10^{-4} M) and then transferred to medium containing different submaximal Ca^{2+} concentrations, there was a graded stepwise release similar to that shown in Fig. 1.

The simplest explanation for the graded vesicular response is that the different vesicles have different threshold sensitivities to Ca^{2+} . However, the same response would be expected for a vesicle population of uniform sensitivity responding at random if the state of each unit (active or inactive) cannot change as long as the Ca^{2+} concentration remains unmodified. A kinetic model that could explain the all or nothing response of the K^+ channel has been proposed by Lew, Muallem and Seymour (see accompanying letter¹³).

Our measurements of the space resealed to ^{86}Rb after vesiculation show that each red cell should provide 200-300 vesicles. Assuming that the number of K^+ channels per cell is less than 200 (ref. 12), the mean frequency of channels per vesicle should be below 1. In agreement with this prediction, we find that a substantial fraction (15-40%) of ^{86}Rb loaded during the vesiculation step is not released by Ca^{2+} , even in the presence of the ionophore A23187, which renders right-side out vesicles sensitive to Ca^{2+} . It thus seems reasonable to assume that most of the Ca^{2+} -sensitive vesicles should have only one or a few channel units. On these bases, the all-or-none nature of the vesicular response does probably reflect the behaviour of individual channel units.

The Ca^{2+} -induced fluxes were specific for Rb (^{22}Na fluxes remained unchanged) and sensitive to quinine (92-97% inhibition at 1 mM) at Ca^{2+} concentrations $\leq 10^{-6}$ M, properties similar to those described in intact cells or resealed ghosts^{3,14-18}. However, at 10^{-4} M Ca^{2+} , we found a quinine-insensitive increase of the fluxes of ^{22}Na (to twice the basal value) and ^{86}Rb (see the difference between the values at pCa 4 and pCa 6 in Fig. 2). This observation suggests a nonspecific effect of Ca^{2+} at such concentrations and is reminiscent of the simultaneous increases of Na^+ and K^+ permeabilities reported pre-

viously in red cells after treatments that presumably increase their intracellular Ca^{2+} content¹⁹.

The study of the kinetics of the Ca^{2+} -induced ^{86}Rb fluxes was complicated by their rapidity ($t_{1/2}$ of ~ 0.5 min) and also by the fact that they could not be described by single exponentials. However, the fraction of the vesicular space equilibrated with the incubation medium at the end of the rapid phase for a given Ca^{2+} concentration was highly reproducible between experiments if expressed as the per cent of the maximal effect. In addition, these values were similar for the uptake and release experiments. The activation of ^{86}Rb fluxes began at 10^{-8} M Ca^{2+} and was almost complete at 10^{-6} M Ca^{2+} , half-maximal activation occurring at a $p\text{Ca}$ of ~ 7.55 (Fig. 2).

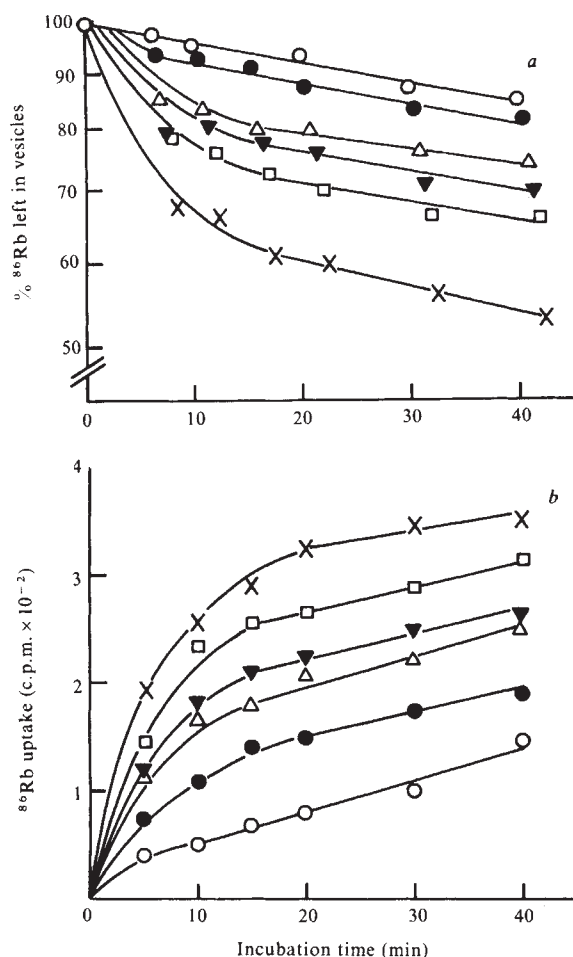


Fig. 1 Effects of external Ca^{2+} on the release (a) and uptake (b) of ^{86}Rb by inside-out vesicles. Vesicles prepared as described by Lew and Seymour¹² were needed, washed once and resuspended in medium of the following composition (mM): KCl, 18; EGTA, 0.1; K-HEPES, 16.5, pH 7.5. For the transport experiments, 0.2-ml aliquots of vesicle suspension (corresponding to ~ 0.2 g of initial cells) were mixed with 0.6–1 ml of medium. ^{86}Rb was added either to the vesiculation medium (release experiments) or to the final incubation medium (uptake experiments). Fixed Ca^{2+} concentrations were obtained using mixtures of EGTA or N-hydroxyethylthylenediamine triacetic acid (HEDTA) (final concentration 0.4 mM) and calcium. Dissociation constants, measured in the incubation medium²⁰, were: EGTA, 4×10^{-8} ; HEDTA, 4×10^{-7} . The incubation period was terminated by running 0.1-ml aliquots of the vesicle suspension through 1-ml columns containing Dowex 50-X8-100 resin^{12,21}. Radioactivity was measured by Cerenkov counting. Symbols represent the following $p\text{Ca}$ values: ≤ 9 (○), 7.85 (●), 7.55 (△), 7.37 (▼), 6.88 (□), 4.00 (×). Vesicular space revealed to ^{86}Rb (ref. 12) ranged over 76–130 μl per g of cells in different experiments. 40–70% of that space was sensitive to Ca^{2+} and this value increased to 60–85% in the presence of 2 μM A23187. Everted vesicular surface⁷ was 25–35% of the total.

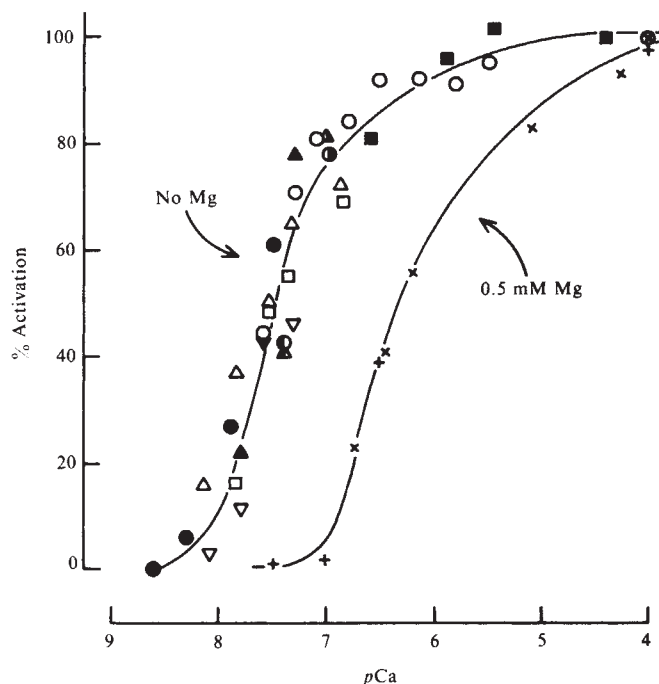


Fig. 2 Activation of ^{86}Rb transport by Ca^{2+} . The vesicular space sensitive to Ca^{2+} (that which equilibrates with the medium during the rapid phase) was calculated from experiments similar to those shown in Fig. 1 by extrapolation to zero-time. Per cent activation was then calculated taking the value obtained without Ca^{2+} as 0% and that obtained at $p\text{Ca}$ 4 as 100%. The results obtained in 14 experiments with vesicles prepared using different blood samples are shown with different symbols. Squares represent data obtained in release experiments while all other symbols correspond to uptake experiments. The curve on the left was obtained in the absence of Mg^{2+} and the one on the right with 0.5 mM Mg^{2+} .

The procedure described here offers an easy way of investigating modifications of the sensitivity to Ca^{2+} of ^{86}Rb transport. The addition of 5×10^{-4} M MgCl_2 to the medium reduced the apparent $p\text{Ca}$ value at which half-maximal activation of ^{86}Rb transport took place ($p\text{Ca}_{1/2}$) from 7.55 to 6.35 (Fig. 2). Smaller Mg^{2+} concentrations (2×10^{-5} , 5×10^{-5} , 10^{-4} , 2×10^{-4} M) produced smaller variations of $p\text{Ca}_{1/2}$ (7.35, 7.15, 6.95, 6.82), a 10-fold increase of Mg^{2+} concentration decreasing $p\text{Ca}_{1/2}$ by 0.58 units. The inhibition of Ca^{2+} -dependent K^+ transport by Mg^{2+} reported here is consistent with previous observations in resealed ghosts¹⁵. The actions of other intracellular effectors on the sensitivity to Ca^{2+} of $\text{K}^+(\text{Rb}^+)$ transport could be investigated similarly. Experiments performed in our laboratory have shown, for example, that neither calmodulin nor NADH modifies the kinetics of activation by Ca^{2+} nor the inhibition by Mg^{2+} or quinine. Work is now in progress to investigate the effects of other intracellular proteins and reducing agents.

We thank Dr V. L. Lew for valuable discussions and J. Alvarez for technical assistance. This study was supported in part by a grant from the Fondo Nacional para el Desarrollo de la Investigación Científica.

Received 3 December 1981; accepted 1 March 1982.

- Meech, R. W. *30th Symp. exp. Biol.*, 161–191 (Cambridge University Press, London, 1976).
- Lew, V. L. & Ferreira, H. G. *Curr. Topics Membranes Transport* **10**, 217–277 (1978).
- Blum, R. M. & Hoffman, J. F. *J. Membrane Biol.* **6**, 315–328 (1971).
- Blum, R. M. & Hoffman, J. F. *Biochem. biophys. Res. Commun.* **46**, 1146–1152 (1972).
- García-Sancho, J., Sanchez, A. & Herreros, B. *Biochim. biophys. Acta* **556**, 118–130 (1979).
- Passow, H. in *The Function of Red Blood Cells: Erythrocyte Pathobiology* (ed. Wallach, D. F. H.) (Liss, New York, 1981).
- Steck, T. L. & Kant, J. A. *Meth. Enzym.* **31**, 172–180 (1974).
- Grimstein, S. & Rothstein, A. *Biochim. biophys. Acta* **508**, 236–245 (1978).
- Lew, V. L., Muallem, S. & Seymour, C. A. *J. Physiol., Lond.* **307**, 36P–37P (1980).
- Sze, H. & Solomon, A. K. *Biochim. biophys. Acta* **554**, 180–194 (1979).
- Sarkadi, B., Szebeni, J. & Gardos, G. in *Membrane Transport in Erythrocytes* (eds Lassen, U. V., Ussing, H. H. & Wieth, J. O.) 220–231 (Munksgaard, Copenhagen, 1980).
- Lew, V. L., Muallem, S. & Seymour, C. A. *J. Physiol., Lond.* **307**, 36P–37P (1980).
- Lew, V. L. & Seymour, C. A. in *Techniques in Lipid and Membrane Biochemistry* (eds Kornberg, H. L. et al.) (North-Holland, Amsterdam, in the press).

13. Lew, V. L., Muallem, S. & Seymour, C. A. *Nature* **296**, 742–744 (1982).
14. Gardos, G. *Biochim. biophys. Acta* **30**, 653–654 (1958).
15. Simons, T. J. B. *J. Physiol., Lond.* **256**, 227–244 (1976).
16. Armando-Hardy, M., Ellory, J. C., Ferreira, H. G., Fleminger, S. & Lew, V. L. *J. Physiol., Lond.* **250**, 32P–33P (1975).
17. Sanchez, A., Garcia-Sancho, J. & Herreros, B. *FEBS Lett.* **110**, 65–68 (1980).
18. Reichstein, E. & Rothstein, A. *J. Membrane Biol.* **59**, 57–63 (1981).
19. Romero, P. J. & Whittam, R. *J. Physiol., Lond.* **214**, 481–507 (1971).
20. Bers, D. M. *J. Physiol., Lond.* **312**, 2P–3P (1981).
21. Gasko, O. D., Knowles, A. F., Shertzer, H. G., Soulinna, E.-M. & Racker, E. *Analyt. Biochem.* **72**, 57–65 (1976).

Intracellular Ca^{2+} activates a fast voltage-sensitive K^+ current in vertebrate sympathetic neurones

P. R. Adams*, A. Constanti†, D. A. Brown† & R. B. Clark

Department of Physiology and Biophysics, University of Texas, Medical Branch, Galveston, Texas 77550, USA

Many neurones, when depolarized, exhibit two components of outward K^+ current—the voltage-sensitive delayed rectifier current originally described in squid axons by Hodgkin and Huxley¹, and an additional current triggered by the entry of Ca^{2+} ions². These two currents have been termed I_K and I_C , respectively³. Previous experiments have indicated that both forms of K^+ current are also present in vertebrate sympathetic neurones^{4–6}. We have now studied the properties of I_C in bullfrog sympathetic neurones, uncontaminated with I_K , by injecting Ca^{2+} ions into the cells and measuring the resultant outward currents under voltage-clamp, in the manner previously used for large molluscan neurones^{7,8}. We find three interesting properties of I_C in these vertebrate neurones. First, it shows strong voltage sensitivity independent of the voltage sensitivity of the Ca^{2+} channels (which have been bypassed by the injection technique). Second, I_C activates and deactivates very rapidly ($\tau_c \leq 20$ ms at 0 mV), with stepped changes in membrane potential. Current fluctuation analysis and patch-clamp records of single-channel currents yielded evidence for appropriate short-lifetime ionic channels with a maximum conductance of ~ 100 pS. Finally I_C in ganglion cells is highly sensitive to external tetraethylammonium. We deduce that in these neurones I_C is a fast current which can contribute a substantial fraction to the repolarizing current during an action potential.

Figure 1a illustrates outward currents evoked in a voltage-clamped bullfrog sympathetic neurone⁹ by brief iontophoretic intracellular injections of Ca^{2+} ions at different holding potentials. The first two panels show currents produced by equal negative and positive iontophoretic pulses at the initial holding potential of -30 mV. There is no change in current-to-ground (upper trace) during the negative iontophoretic pulse because the iontophoretic current was balanced by positive clamp current and no additional active membrane current was induced. However, during positive iontophoresis an unbalanced outward membrane current appears, outlasting the iontophoretic injection. With such brief (< 1 s) iontophoretic applications, repeatable currents could be recorded at 5–10-s intervals without decrement. At any given holding potential the peak outward current increased linearly with increasing iontophoretic currents up to 10 nC. However, the sensitivity varied greatly from cell to cell: it was usually within the range 1–3 nA outward current per nC injected current at -30 mV, but some cells showed no response to 10 nC or more. Assuming a transport number of 0.3 for Ca^{2+} in 0.2 M CaCl_2 solution⁷, 10 nC current should expel $\sim 1.5 \times 10^{-14}$ mol of Ca^{2+} ion; in a spherical cell of

~ 40 μm diameter this is equivalent to an increment of ~ 0.5 mM internal $[\text{Ca}^{2+}]$, although the increase in free ionized $[\text{Ca}^{2+}]$ is presumably much less because of buffering and redistribution. Unlike the situation in large molluscan neurones⁷, vertical movement of the iontophoretic pipette did not greatly alter the magnitude or time course of the current response, suggesting that in the smaller ganglion cells the intracellular Ca^{2+} concentration very rapidly attained homogeneity. Control injections to ± 10 nC from electrodes filled with 0.3 M KCl at pH 5.8 did not produce outward membrane currents.

The outward current generated by constant iontophoretic Ca^{2+} injections increased steeply as the membrane was depolarized between -50 and -20 mV, as shown on the right of Fig. 1a. In normal Ringer solution the current fell to zero around -50 mV and did not reverse with further hyperpolarization. Clear reversal to an inward current (at -35 ± 2 mV ($\pm$ s.e.m.); $n = 8$) was obtained on raising $[\text{K}^+]_{\text{out}}$ to 25 mM. Assuming the underlying conductance change G_c at any membrane potential to be unaffected by raising $[\text{K}^+]_{\text{out}}$, the hypothetical reversal potential in normal Ringer solution (containing 2.5 mM K^+) could be calculated from the ratio of the currents evoked at a fixed holding potential in normal and high $[\text{K}^+]$ solutions and from the reversal potential in high $[\text{K}^+]$. The mean extrapolated reversal potential in normal Ringer was estimated at -78 ± 7 mV ($n = 5$), and the positive shift on raising $[\text{K}^+]$ 10-fold was $+41 \pm 7$ mV.

At a reversal potential of -78 mV, the voltage dependence of I_C shown in Fig. 1a accorded with an exponential increase in the underlying conductance (ΔG_c) for each 7.5 mV depolarization above -50 mV. In seven experiments of this type, the average slope was 11 mV per e -fold ΔG_c up to -10 mV. This is appreciably steeper than that previously reported (25 mV (ref. 8) and 28 mV (ref. 10)) for the voltage-sensitivity of I_C activated by intracellular Ca^{2+} injections in molluscan neurones. Removal of external Ca^{2+} or addition of the Ca^{2+} -channel blocker Cd^{2+} in concentrations (0.2–0.5 mM) sufficient to block the inward Ca^{2+} current in bullfrog sympathetic neurones¹¹ did not materially change this voltage sensitivity, showing that it was related solely to the action of internal Ca^{2+} and not to any inward Ca^{2+} current. Steady outward currents at depolarized potentials were reduced in Cd^{2+} solution, which allowed the membrane potential to be held at more positive potentials than in normal Ringer solution. This revealed that, at potentials positive to -10 mV, I_C tended to saturate and, in some cells, actually diminished with increasing depolarization. This could not be alleviated by reducing the iontophoretic Ca^{2+} current.

When a voltage step was applied during a Ca^{2+} injection, the outward current showed an exponential growth or decline with time, and a time constant of ~ 15 ms at -10 mV command potential (Fig. 2a). To study these relaxations in more detail, the single microelectrode clamp method (which had limited current-passing capacity) was replaced with a conventional two-electrode clamp⁹ and I_C was generated either by Ca^{2+} injection or by activating a sustained inward Ca^{2+} current with a long depolarizing command. Figure 3 shows examples of such depolarizing commands, in another context. During sustained depolarization the delayed rectifier current becomes totally inactivated so that, if no Ca^{2+} is present in the external solution, only a leak current is recorded and a superimposed voltage step produces a square 'ohmic' current response (Fig. 2b, left-hand record). When Ca^{2+} is added to the external solution the inward Ca^{2+} current generates an additional outward I_C (which is blocked by Cd^{2+}) and a voltage step now produces a time-dependent relaxation comparable with that seen during a Ca^{2+} injection (Fig. 2b, right panel). The average time constant for this relaxation (τ_c) was 16.0 ± 0.9 ms ($n = 10$) at -10 mV. Over the range -80 to $+30$ mV, τ_c showed an exponential lengthening for each 27 mV depolarization. An equivalent voltage sensitivity of these relaxations was confirmed when Ca^{2+} was injected through a third intracellular microelectrode.

The current record during long depolarizations in Ca^{2+} solution was distinctly 'noisy'. This noise disappeared on removing

* Present address: Department of Neurobiology, State University of New York, Stony Brook, New York 11790, USA.

† Permanent address: Department of Pharmacology, The School of Pharmacy, University of London, 29/39 Brunswick Square, London WC1N 1AX, UK.

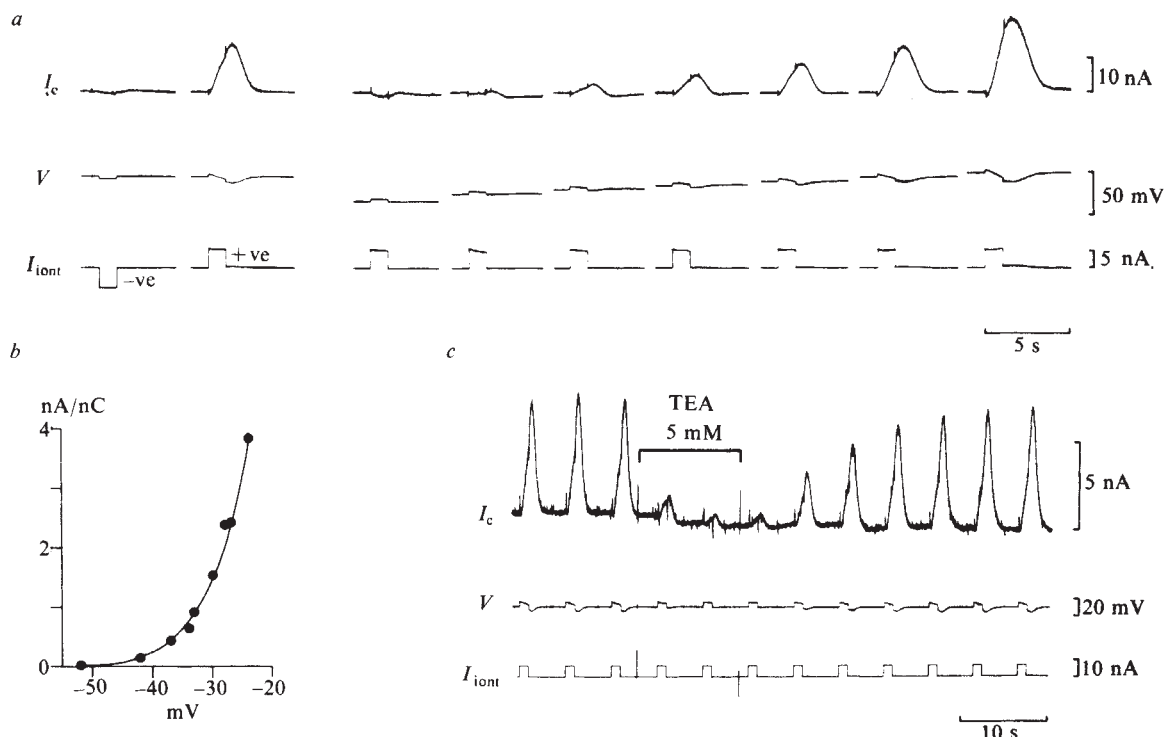


Fig. 1 Outward currents generated by intracellular Ca^{2+} injections in voltage-clamped bullfrog sympathetic neurones. Lumbar sympathetic neurones were maintained at 22 °C in normal Ringer solution containing 2.5 mM KCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$ and 2 mM Tris buffered to pH 7.2 with HCl. Neurones were impaled with two electrodes drawn from thin-walled glass tubing (tip resistances 8–20 M Ω). One electrode, filled with 0.2 or 0.4 M $CaCl_2$ solution in distilled water (pH 5.8), was used for iontophoretic injection of Ca^{2+} . The other electrode, filled with 3 M KCl, was used to record membrane potential and to inject clamp current via a sample-and-hold voltage-clamp preamplifier (Dagan 8100, switching between voltage sampling and current injection at 3 kHz). Clamp current (upper trace in a and c, I_c) was recorded as true current-to-ground, filtered at 300 Hz; iontophoretic current (lower trace in a and c, I_{iont}) was recorded by subtracting Dagan current (continuous output) from current-to-ground. The records in a show responses to 1 s, 6 nA iontophoretic injections of Ca^{2+} at different membrane holding potentials between –55 and –20 mV, preceded by test responses to negative and positive iontophoretic currents at –30 mV holding potential. b, Shows the change in peak outward current with holding potential. The potential on the abscissa is the recorded membrane potential at the height of the current response. The smooth curve shows how the current would change if the underlying conductance increased e-fold for 7.5 mV depolarization and if the current reversal potential were –78 mV. c, Shows the effect of tetraethylammonium (TEA) bromide on the Ca^{2+} -generated currents in another cell at a holding potential of –30 mV.

external Ca^{2+} or adding Cd^{2+} . Spectral analysis of the Ca^{2+} -dependent noise during the latter stages of the depolarization (when the a.c. current record was flat) yielded Lorentzian power spectra (Fig. 3b); spectra obtained at rest, during hyperpolarizing commands or during depolarizations in the absence of Ca^{2+} or presence of Cd^{2+} , always resembled the background spectrum shown in Fig. 3b. The average corner frequency (f_c) of the Lorentzian spectra at a command potential of –10 mV was 7.9 ± 1.5 Hz (mean $\pm$ s.e.m.; $n = 6$). This corresponds to a time constant ($\tau = 1/2\pi f_c$) of 20 ms (16.9–24.9 ms), which is similar to that for the current relaxation following a voltage jump to –10 mV. Assuming the fluctuations to arise from transient openings and closings of many individual ionic channels, each with a fixed conductance (see below), the average single-channel current can be estimated from the ratio of the total variance to the mean current¹². Total variance was calculated from the area under the test spectrum between 0.1 and 200 Hz, after subtraction of the background spectrum. The average value of the macroscopic current was taken as the median value during the spectral computation—a reasonable approximation for long depolarizations and near-linearly decaying currents. Mean I_c was determined by subtracting the current during equal but opposite hyperpolarizations, or during depolarization in the presence of Cd^{2+} or absence of Ca^{2+} . (Residual currents were rather small, and imprecision in the measurement of mean I_c contributed greatly to the scatter on the resultant ratios.) In one cell clamped to 0, +10 and +20 mV, the variance increased linearly with the mean current. This suggests that the probability of channel opening was relatively low, so that the non-stationary currents would not introduce significant distortion into the spectra if analysis is restricted to the period when the a.c. current record is flat¹³, as in our experiments. The mean ratio of variance to mean I_c was 8.1 ± 1.7 pA (mean $\pm$ s.e.m.; $n = 17$ cells), giving

an average single-channel conductance of $\sim 100 \pm 20$ pS at a reversal potential of –78 mV.

In ganglion cells cultured for 1–2 months *in vitro*¹⁴ (which conveniently had naked cell membrane), appropriately large single channels could be detected using the gigohm-seal patch-clamp technique^{15,16}. An example is shown in Fig. 4. At rest potential (around –60 mV in the unimpaled cells, judging from the better values obtained on impalement) the current record was quiescent. When the electrode was polarized, thus depolarizing the enclosed membrane patch, one or more discrete current channels appeared whose opening frequency and duration clearly increased with increasing depolarization. During weak depolarization the channels showed simple open-closed behaviour, but at depolarization ≥ 80 mV the open state was interrupted by brief closures. These became more frequent with increasing positivity, giving a very noisy open-state current level. Up to 100 mV depolarization the single-channel current increased linearly with applied potential, with a slope corresponding to ~ 100 pS open-channel conductance. Similar values were obtained in two other patches. However, at more positive potentials this linearity was lost and at ≥ 175 mV the peak single-channel current actually declined: such behaviour was seen in one other patch which could be polarized sufficiently. At 60 mV patch potential (~ 0 mV membrane potential) the mean channel open time was ~ 20 ms. The apparent open time increased exponentially per 22 mV depolarization up to +150 mV patch potential.

These single-channel events are clearly similar to those recently described in isolated mammalian chromaffin cells¹⁷ and cultured embryonic muscle cells¹⁸, in terms of both the large single-channel conductances and the open state flutterings. The latter may result from superimposed channel blocking¹⁶ or might relate to the vibrations between open and closed states

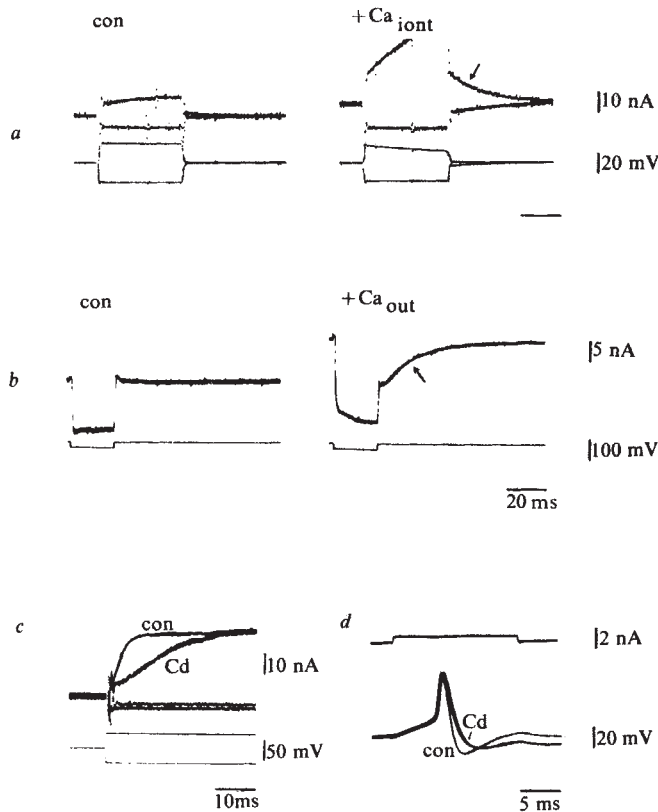


Fig. 2 Kinetic tests on I_C . *a*, Voltage jumps of +10 mV and -10 mV applied from a holding potential of -10 mV before (con) and during (+ Ca_{iont}) intracellular iontophoresis of 1 nA Ca. The single microelectrode voltage clamp technique referred to in Fig. 1 was used to record voltage and pass current: the limited current-passing capacity leads to some clamp failure during the large outward current but not during the repolarizing tail current (arrowed). *b*, 30 mV hyperpolarizing voltage jumps applied from a steady prepotential of -10 mV in Ca^{2+} -free solution (con) and in 2 mM Ca^{2+} solution (+ Ca_{out}). Two-electrode voltage clamp, no Ca^{2+} injection. Note the relaxation (arrowed) in the presence of Ca^{2+} . *c*, Voltage commands of ± 40 mV applied from a holding potential of -50 mV in normal Ringer solution (con) and then in Ca^{2+} -free solution containing 0.5 mM Cd^{2+} (Cd). Two-electrode clamp. *d*, Action potentials generated by depolarizing current injections in normal Ringer solution (con) and in Ringer containing 0.5 mM Cd^{2+} (Cd), recorded using independent voltage-recording and current-passing electrodes.

seen in *n*-acetylglutamic channels¹⁹. As in the chromaffin cell experiments¹⁵, it was not necessary to have Ca^{2+} in the pipette to observe channels, so presumably there was sufficient free Ca^{2+} inside the cell to trigger them. Single I_C channels have also been detected in molluscan neurones²⁰, of conductance (19 pS) rather lower than but mean open times comparable with those in the ganglion cells, and in neuroblastoma cells (C. F. Stevens, personal communication).

As the cultured cells showed macroscopic C-currents of similar amplitude, time constant and voltage sensitivity to those in freshly dissected ganglia, we suppose that the single channels represent the subunits of these macroscopic currents. If this is so, the comparable duration and voltage sensitivities of the channel open times and relaxation times suggest that the latter are determined primarily by the channel closing rate. This implies that a relatively small proportion of the available channels was activated by the intracellular Ca^{2+} injections or by the imposed depolarizations. The fall in single-channel conductance at very positive potentials might explain the saturation of the macroscopic current with increasing depolarization. However, we also detected signs of a loss of voltage sensitivity at high internal Ca^{2+} concentrations. This was predicted from the observations on chromaffin cell channels¹⁷, where large or repeated Ca^{2+} injections abolished the relaxations previously generated by voltage jumps. The origin of the voltage sensitivity is not clear: voltage dependence seems too steep to be accounted for entirely by voltage-sensitive binding of Ca^{2+} to a site within the membrane potential field⁸, unless the binding of more than one Ca^{2+} ion per channel was necessary²¹ or an additional energy barrier were imposed²²; alternatively, voltage sensitivity might be due to the gating of the K^+ channels themselves.

Two further points of interest concerning ganglionic C-channels relate to their pharmacology and physiological function. First, like some^{23,24} but not all^{3,25} molluscan neurones, I_C in these ganglion cells was at least as sensitive as the delayed rectifier current to external tetraethylammonium (TEA). Thus, the currents generated by Ca^{2+} injections (Fig. 1c) or by membrane depolarization were virtually abolished by 5 mM TEA, and were half-reduced at 1 mM. In agreement with this, TEA suppressed both voltage-jump current relaxations and current noise. Thus, the effect of TEA on these neurones cannot be attributed unequivocally to inhibition of delayed rectifier currents. In contrast, I_C was insensitive to 1 mM external 4-aminopyridine.

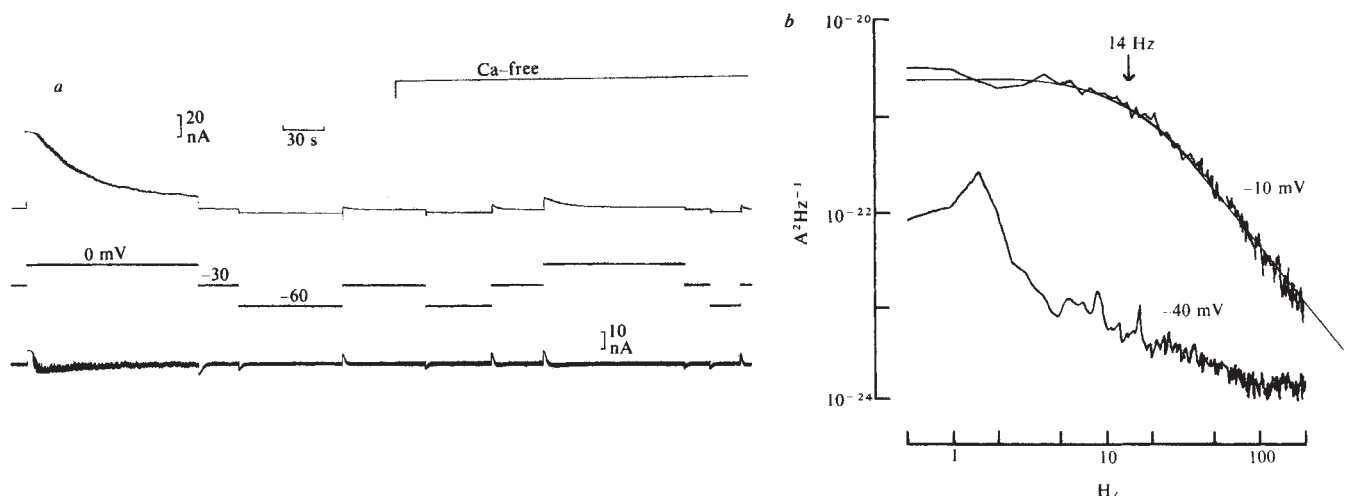
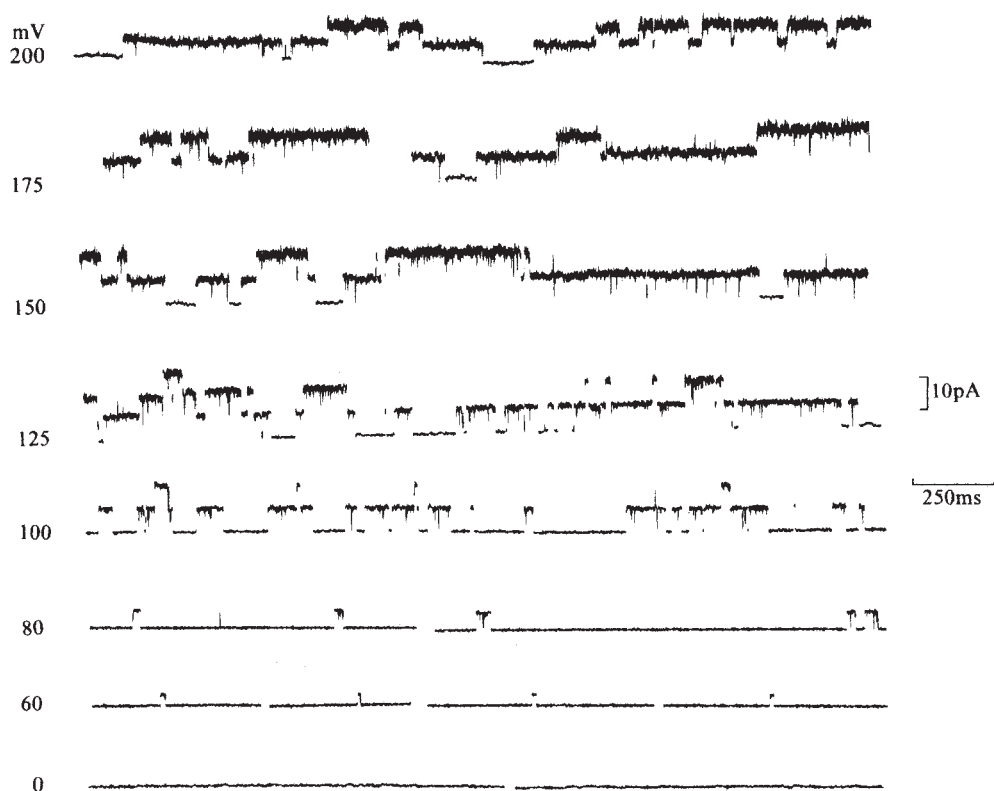


Fig. 3 Current noise associated with I_C recorded from freshly dissected ganglion cells with a two-electrode voltage clamp. The records in *a* show currents evoked by long ± 30 -mV commands from a holding potential of -30 mV in normal Ringer solution and after replacement with Ca^{2+} -free Ringer. Upper trace, d.c. current; middle trace, voltage; lower trace, a.c. current (0.1 Hz high-pass filter). Note that, in the absence of Ca^{2+} , the delayed rectifier current evoked by the depolarizing command subsided to a level no greater than the leak current produced by the preceding hyperpolarizing command. In the presence of Ca^{2+} there is an additional outward current showing appreciable noise. *b*, Power spectra of current noise (sum of 32 spectra) in another cell recorded at the holding potential (-40 mV) and during the last 64 s of a 2-min depolarizing command to -10 mV in normal Ringer solution (corrected by subtraction of the -40 mV spectrum, although this made negligible difference). The curve is drawn according to the equation $S(f) = S(0)/[1 + (f/f_c)^2]$ where $S(f)$ is the power ($\text{A}^2 \text{Hz}^{-1}$, ordinates) at frequency f (Hz, abscissa), $S(0)$ is the power at $f = 0$, and f_c is 14 Hz. The mean leak-corrected I_C during spectral analysis was 9.4 nA.

Fig. 4 Single-channel activity in a cultured bullfrog sympathetic ganglion cell. A fire-polished pipette filled with normal Ringer solution (tip resistance 3–5 M Ω) was lowered vertically on to the cell membrane and suction gently applied to produce a gigaohm seal. The enclosed membrane patch was initially silent but after a few minutes the electrical events illustrated in these records appeared and lasted for ~1 h. The voltages on the left show the potential of the pipette with respect to ground (thus, 0 means that the enclosed membrane patch was at rest potential; a patch potential of 60 means that the membrane patch was depolarized by 60 mV). Recording bandwidth, ~1 kHz.



The second point is that, during a depolarizing voltage step following a Ca^{2+} injection, I_C develops exponentially with a time constant at least as short as that of the delayed rectifier current (which is sigmoidal and relatively slow in these neurones²⁶). Because the inward Ca^{2+} current in these cells is also fast ($\tau_{\text{max}} \sim 4$ ms at 0 mV)¹¹, this suggests that I_C might be activated relatively early during a depolarizing event. Figure 2c shows that this is indeed the case, as addition of Ca^{2+} and/or removal of external Ca^{2+} slowed outward current development between -20 and +30 mV. Furthermore, in addition to reducing the spike after-hyperpolarization^{4,5}, Cd^{2+} clearly slowed spike repolarization (Fig. 2d). Thus, in at least a proportion of bullfrog sympathetic neurones, the K^+ current triggered by rapid Ca^{2+} entry may subserve one of the functions more normally fulfilled by the delayed rectifier current.

This work was supported by NIH grants NS 14986 and NS 14920 to P.R.A., an MRC grant to D.A.B., a Wellcome Travel Grant to A.C. and an MDA Fellowship to R.B.C. We thank Claire Adams for culturing the neurones.

Received 14 September 1981; accepted 17 February 1982.

- Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **116**, 449–472 (1952a).
- Meech, R. W. A. *Rev. Biophys. Bioengng* **7**, 1–18 (1978).
- Thompson, S. H. *J. Physiol., Lond.* **265**, 465–488 (1977).
- Minota, S. *Jap. J. Physiol.* **24**, 501–512 (1974).
- McAfee, D. A. & Yarowski, P. J. *J. Physiol., Lond.* **290**, 507–523 (1979).
- Adams, P. R., Constanti, A. & Brown, D. A. *Fedn Proc.* **39**, 2063 (1980).
- Gorman, A. L. F. & Hermann, A. *J. Physiol., Lond.* **296**, 393–410 (1979).
- Gorman, A. L. F. & Thomas, M. V. *J. Physiol., Lond.* **308**, 287–313 (1980).
- Brown, D. A. & Adams, P. R. *Nature* **283**, 673–676 (1980).
- Hermann, A. & Hartung, K. *Pflügers Arch. ges. Physiol.* **389**, R22 (1981).
- Adams, P. R. *Adv. Physiol. Sci.* **4**, 135–138 (1981).
- Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655–692 (1973).
- Sigworth, F. J. *Biophys. J.* **35**, 289–300 (1981).
- Siggins, G. R., Gruol, D. L., Padjen, A. L. & Forman, D. S. *Nature* **270**, 263–265 (1977).
- Horn, R. & Patlak, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6930–6934 (1980).
- Hamill, O., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Marty, A. *Nature* **291**, 497–500 (1981).
- Pallotta, B. S., Magleby, K. L. & Barrett, J. N. *Nature* **293**, 471–474 (1981).
- Szabo, G. *Biophys. J.* **33**, 64a (1981).
- Lux, H. D., Neher, E. & Marty, A. *Pflügers Arch. ges. Physiol.* **389**, 293–295 (1981).
- Meech, R. W. & Thomas, R. C. *J. Physiol., Lond.* **298**, 111–129 (1980).
- Lux, H. D. *Molluscan Nerve Cells* (eds Koester, J. & Byrne, J. H.) 105–114 (Cold Spring Harbor Laboratory, New York, 1980).
- Meech, R. W. & Standen, N. B. *J. Physiol., Lond.* **249**, 211–239 (1975).
- Hermann, A. & Gorman, A. L. F. *J. gen. Physiol.* **78**, 87–110 (1981).
- Kostyuk, P. G., Doroshenko, P. A. & Tsyndrenko, A. *Neuroscience* **5**, 2187–2192 (1980).
- Adams, P. R., Brown, D. A. & Constanti, A. *J. Physiol.* (in the press).

Vasopressin excites hippocampal neurones

M. Mühlethaler & J. J. Dreifuss

Department of Physiology, University of Geneva Medical School, 1211 Geneva 4, Switzerland

B. H. Gähwiler

Preclinical Research, Sandoz Ltd, 4002 Basle, Switzerland

Hypothalamic supraoptic and paraventricular neurones were the first peptidergic cells to be described in the mammalian brain¹. Most of these large nerve cells synthesize either oxytocin or vasopressin², as well as their respective neurophysins, and project to the posterior lobe of the pituitary gland. This system of neurones is responsible for the release of these peptides into blood vessels of the posterior lobe in response to appropriate stimuli. Data obtained using immunoassay and immunocytochemical methods, however, have revealed vasopressin-like immunoreactive material^{3,4} as well as vasopressin-positive axons^{5–7} and presynaptic terminals⁸ in other central nervous system (CNS) locations. Moreover, vasopressin has been shown to affect some aspects of animal behaviour, including memory retention^{9–13}. It has therefore been suggested that vasopressin may act as a neurotransmitter or neuromodulator at synapses in the brain. We show here that low concentrations (10^{-8} – 10^{-6} M) of vasopressin powerfully and reversibly increase the rate of firing of neurones in the CA1 area of hippocampal slices from rat and that this effect can be fully antagonized by an anti-vasopressor vasopressin analogue. Hippocampal neurones obtained from Brattleboro rats were also excited by exogenous vasopressin.

We prepared transverse slices from the ventral hippocampus of male rats (~250 g body weight), as vasopressin-positive nerve fibres detected by immunocytochemical methods have been shown to terminate on the dendritic tree of pyramidal neurones

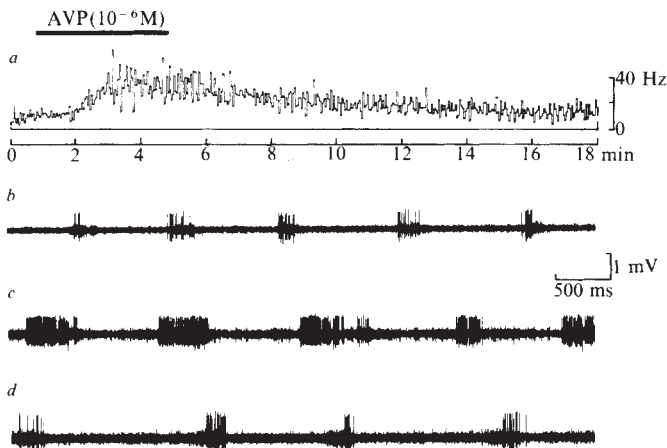


Fig. 1 Effect of vasopressin (AVP, 10^{-6} M) on the rate of firing of a CA1 neurone in a hippocampal slice. *a*, Ratemeter record; AVP was added to the perfusion medium during the period indicated by the thick bar. *b–d*, Oscilloscope tracings showing extracellular records (electrode filled with NaCl, 4 M; tip resistance, 5–10 M Ω) of the activity of the same cell before (*b*), during (*c*) and after (*d*) application of AVP.

in the ventral hippocampus^{5,6,14}. Slices (~ 500 μ m thick), were laid on a nylon grid at the interface of a perfusion solution (composition (mM): NaCl 124.0; KCl 5.0; NaHCO₃ 26.0; MgSO₄ 1.4; CaCl₂ 1.0; KH₂PO₄ 1.24; glucose 10.0; temperature 37°C) and an humidified, oxygenated (95% O₂, 5% CO₂) gas mixture. A glass micropipette was lowered under visual guidance into the pyramidal cell layer in the CA1 area of the hippocampus. Substances tested included arginine-vasopressin (AVP; Bachem or Vega), lysine-vasopressin (LVP; Bachem, Sandoz or Serva), 1-desamino-D-arginine-vasopressin (DDAVP; Ferring), desglycinamide-arginine-vasopressin (DGAVP; Organon), oxytocin (Bachem) and the vasopressor antagonist [1-(β -mercapto- β , β -cyclopentamethylpropionic acid)2-(*O*-methyl)tyrosine]arginine-vasopressin (Peninsula). These compounds were dissolved in the bathing solution immediately before application and perfused at a rate of 2.0 ml min⁻¹. Extracellular recordings were obtained in 43 slices from cells which displayed short intermittent bursts of activity in control solution.

AVP and LVP induced a sustained increase in firing rate in 56 of 63 of these cells and prolonged the duration of the bursts without markedly altering the inter-burst intervals (Fig. 1). The threshold concentration of vasopressin which accelerated cell firing was $\sim 10^{-8}$ M (Fig. 2). Maximal stimulation with AVP occurred at $\sim 10^{-6}$ M and resulted in an increase in mean firing rate from 13.5 ± 1.1 to 44.0 ± 3.2 spikes per s (mean $\pm$ s.e.m., $n = 47$). These concentrations of vasopressin are much lower than those required for the stimulant action of acetylcholine or glutamate ($\sim 10^{-3}$ M), but similar to those at which enkephalin excited hippocampal pyramidal neurones^{15,16}.

The stimulatory effect of vasopressin was slowly reversible and little or no tachyphylaxis was apparent on repeated vasopressin application. We therefore attempted to compare the magnitude of the response with that produced by natural or synthetic analogues. The other mammalian neurohypophysial peptide, oxytocin, which differs from AVP and LVP in amino acid positions 3 and 8 of the molecule, excited hippocampal neurones as effectively as vasopressin in 5 of 6 cells. In contrast, DGAVP, which lacks the glycine terminal present in position 9, failed to excite hippocampal neurones ($n = 6$), even when applied at 10^{-5} M, and glycineamide (5×10^{-4} M) alone had no effect on cell firing ($n = 3$). DGAVP essentially lacks endocrine activity, but has been reported to exert an effect similar to that of vasopressin on memory and learning¹⁷.

Brattleboro rats, which are incapable of vasopressin biosynthesis, show behavioural deficits which can be alleviated by injecting vasopressin¹⁸. To determine whether these rats possess

in their CNS the same binding sites for neurohypophysial peptides as normal non-diabetic rats, we compared the effects of exogenous vasopressin on cells in hippocampal slices obtained from Long-Evans rats ($n = 8$) and from homozygous Brattleboro rats ($n = 7$). In the latter, 10^{-7} M vasopressin accelerated neuronal firing and 10^{-6} M did so more powerfully. Dose-response relationships did not differ significantly between the two strains. As in normal rats, the stimulatory effect of vasopressin was reversibly abolished by the synthetic vasopressin antagonist.

To determine the site of action of vasopressin, small regions (~ 1 mm² surface area) of the CA1 area were cut from hippocampal slices. Vasopressin was as effective in these trimmed slices as in normal slices. Two further sets of experiments suggested that vasopressin can act postsynaptically. First, vasopressin was applied to slices in which the neurones had been synaptically uncoupled by raising the external Mg²⁺ concentration to 10 mM. The cells were no longer spontaneously active, but in 4 of 6 cells, they were still excited by 10^{-6} M AVP. Second, the stimulatory effect of vasopressin was not reduced by applying antagonists to various suspected transmitters. Thus neither atropine, nor propranolol nor naloxone, interfered with the action of vasopressin.

At least two types of peripheral vasopressin receptor have been recognized. The renal receptor (V₂-receptor) is responsible for the antidiuresis induced by vasopressin and is linked to an adenylate cyclase^{19,20}. The V₁-receptor, which is not coupled to an adenylate cyclase, is present on smooth muscle cells and liver cells^{21,22} and triggers the vasopressor and glycogenolytic effects of vasopressin on these tissues. The availability of rather selective V₂-agonists allowed the testing of the nature of the hippocampal response. DDAVP has virtually no effect on the V₁-receptor but is a powerful antidiuretic compound; in contrast, AVP and LVP exert both effects equally²³. In all four cells examined, DDAVP had a much weaker action on hippocampal cell firing than the same concentration of either AVP or LVP. Further evidence that the effects of vasopressin on hippocampal neurones were unlikely to be mediated by V₂-like receptors was obtained from application of a V₁-antagonist²⁴. This compound alone, at 10^{-6} M, had no effect, but

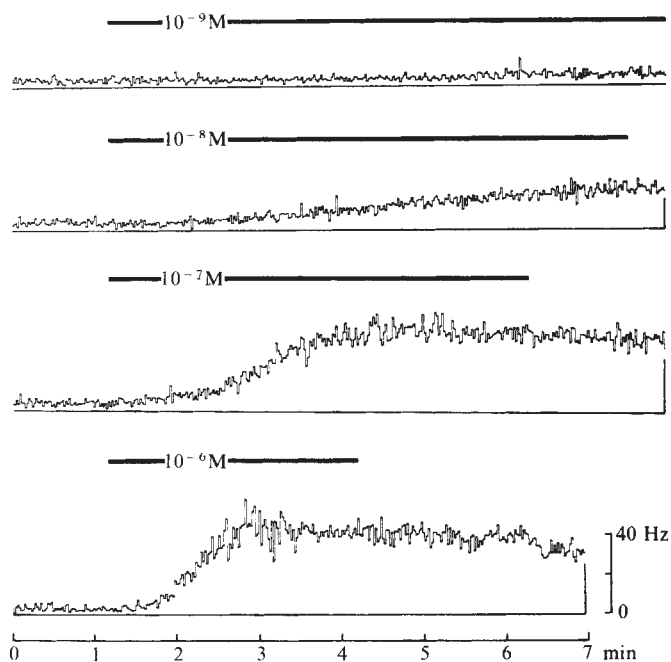


Fig. 2 Segments of ratemeter records illustrating the effects of increasing concentrations of AVP applied to the bath on the rate of firing of a hippocampal neurone, recorded extracellularly. Duration of application (solid black line) and molar concentrations are indicated above each trace.

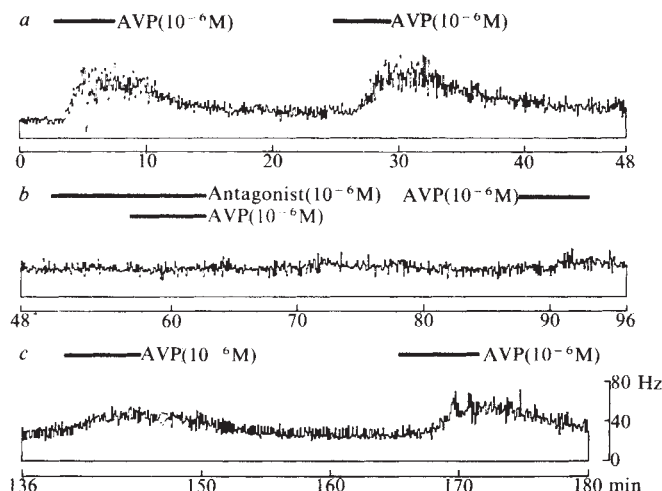


Fig. 3 Effect of the vasopressor antagonistic AVP analogue on the spontaneous and AVP-induced firing of a CA1 hippocampal neuron. Extracellular recording. AVP (10^{-6} M), applied at intervals of ~25 min, caused a marked increase in the rate of firing of this cell without showing any evidence of tachyphylaxis (a). The antagonist (10^{-6} M) did not affect firing *per se*, but blocked the effect of AVP for a prolonged period (b); its effect was rapid in onset, but slowly reversible (c). Note that 40 min of record have been deleted between b and c. Peptides were applied during the periods indicated by the thick black lines.

totally and reversibly abolished the effect of vasopressin on cell firing in all five cells tested (Fig. 3).

In invertebrate ganglia, vasopressin (10^{-9} – 10^{-7} M) induced or prolonged the burst of pacemaker neurones^{25,26}. In the rat brain, iontophoretically applied vasopressin decreased the rate of firing of cells recorded from the hypothalamic supraoptic nucleus²⁷ and increased neural firing in the locus coeruleus²⁸. In the present study, AVP and certain related peptides have been shown to increase the firing rate of hippocampal neurones when added to the medium at concentrations $>10^{-9}$ M. This threshold concentration is at least 50 times higher than the concentration of vasopressin found in cerebrospinal fluid (CSF)²⁹. Even though some peptide degradation by proteolysis may have occurred in the slices and the real concentration at the site where the effects of vasopressin arose is therefore unknown, it seems unlikely that the effect of vasopressin in the hippocampus *in situ* is mediated by the CSF. We interpret the results as evidence that vasopressin can act as a transmitter or modulator in the mammalian hippocampus. Indeed, other investigators have reported that cells of the CA1 region of the hippocampus are stimulated by several peptides, including enkephalins^{15,16} and somatostatin, vasoactive intestinal peptide, bombesin and CCK-8S (refs 30, 31).

Numerous data have suggested an involvement of vasopressin in facilitation of memory storage and retrieval^{9–13}. DGAVP has similar activity whereas oxytocin has the opposite effect. Here we report a different structure–activity relationship for the action of vasopressin in hippocampal slices: both vasopressin and oxytocin are powerful excitants of neural firing, while comparable concentrations of DGAVP have no effect.

This work was supported in part by grant 3.469.79 from the Swiss NSF.

Received 16 December 1981; accepted 12 March 1982.

- Bargmann, W. & Scharrer, E. *Am. Scient.* **39**, 255–259 (1951).
- Vandesande, F. & Dierckx, K. *Cell Tissue Res.* **164**, 153–162 (1975).
- Dogterom, J., Snijdwint, F. G. M. & Buijs, R. M. *Neurosci. Lett.* **9**, 341–346 (1978).
- Glick, S. & Brownstein, M. J. *Life Sci.* **27**, 1103–1110 (1980).
- Buijs, R. M., Swaab, D. F., Dogterom, J. & Van Leeuwen, F. W. *Cell Tissue Res.* **186**, 423–433 (1978).
- Buijs, R. M. *Cell Tissue Res.* **192**, 423–435 (1978).
- Sofroniew, M. V. *J. Histochem. Cytochem.* **28**, 475–478 (1980).
- Buijs, R. M. & Swaab, D. F. *Cell Tissue Res.* **204**, 353–365 (1979).
- De Wied, D. *Life Sci.* **19**, 685–690 (1976).
- Bohus, B., Urban, L., van Wimersma Greidanus, T. B. & De Wied, D. *Neuropharmacology* **17**, 239–247 (1978).

- Kovacs, G. L., Bohus, B., Versteeg, D. H. G., de Kloet, E. R. & De Wied, D. *Brain Res.* **175**, 303–314 (1979).
- De Wied, D. & Versteeg, D. H. G. *Fedn Proc.* **38**, 2348–2354 (1979).
- De Wied, D. *Proc. R. Soc. B210*, 183–195 (1980).
- Buijs, R. M. *J. Histochem. Cytochem.* **28**, 357–360 (1980).
- Lynch, G. S., Jensen, R. A., McGaugh, J. L., Davila, K. & Oliver, M. *Expl Neurol.* **71**, 527–540 (1981).
- Gähwiler, B. H. & Herrling, P. L. *Regul. Peptides* **1**, 317–326 (1981).
- De Wied, D., Greven, H. M., Lande, S. & Witter, A. *Br. J. Pharmac.* **45**, 118–122 (1972).
- De Wied, D., Bohus, B. & van Wimersma Greidanus, T. B. *Brain Res.* **85**, 152–156 (1975).
- Rajerson, R., Marchetti, J., Roy, C., Bockaert, J. & Jard, S. *J. biol. Chem.* **249**, 6390–6400 (1974).
- Roy, C., Barth, T. & Jard, S. *J. biol. Chem.* **250**, 3149–3156 (1975).
- Kirk, C. J., Rodriguez, L. M. & Hems, D. A. *Biochem. J.* **178**, 493–496 (1979).
- Kirk, C. J., Michell, R. H. & Hems, D. A. *Biochem. J.* **194**, 155–165 (1981).
- Sawyer, W. H., Grzonka, Z. & Manning, M. *Molec. cell. Endocr.* **22**, 117–134 (1981).
- Kruszinski, M. *et al. J. med. Chem.* **23**, 363–368 (1980).
- Barker, J. L. & Gainer, H. *Science* **184**, 1371–1373 (1974).
- Barker, J. L., Ifshin, M. & Gainer, H. *Brain Res.* **84**, 501–503 (1975).
- Nicoll, R. A. & Barker, J. L. *Brain Res.* **35**, 501–511 (1971).
- Olpe, H. R. & Baltzer, V. *Eur. J. Pharmac.* **77**, 377–378 (1981).
- Mens, W. B. J., Bouman, H. J., Bakker, E. A. D. & van Wimersma Greidanus, T. B. *Eur. J. Pharmac.* **68**, 89–92 (1980).
- Dodd, J. & Kelly, J. S. *Nature* **273**, 674–675 (1978); *Brain Res.* **205**, 337–350 (1981).
- Dodd, J., Kelly, J. S. & Said, S. I. *Br. J. Pharmac.* **66**, 125 P (1979).

Control of cyclic nucleotide metabolism by non-cholinergic, non-adrenergic nerves in rat thyroid gland

S. T. Green*, J. Singh† & O. H. Petersen†‡

* Department of Physiology, University of Dundee, Dundee DD1 4HN, UK

† The Physiological Laboratory, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK

In addition to the well-established role of thyrotropin (TSH) in controlling metabolism and secretion of thyroid follicular cells^{1–3}, there is evidence to suggest nervous (cholinergic and adrenergic) control mechanisms^{4–8}. Nerves showing immunoreactivity to vasoactive intestinal peptide (VIP) have recently been demonstrated in the thyroid gland⁹, while exogenous VIP has been shown to elevate tissue cyclic AMP levels, increase the number of colloid droplets in follicular cells and increase thyroid hormone secretion *in vivo*. It was therefore suggested that VIP-containing nerves might control thyroid activity⁹, although a functional non-cholinergic, non-adrenergic innervation of the thyroid has not yet been demonstrated. By adopting the technique of electrical field stimulation previously used on isolated pancreatic segments^{10–13}, we have studied cyclic nucleotide metabolism in isolated thyroid fragments and have now demonstrated directly the influence of both cholinergic and adrenergic as well as non-cholinergic, non-adrenergic nerves on thyroid cyclic AMP and cyclic GMP metabolism. In the presence of complete cholinergic and adrenergic blockade, nerve stimulation evoked complex time course changes in tissue cyclic AMP and cyclic GMP levels which could be mimicked precisely by VIP. Electrical stimulation of intrinsic thyroid nerves released, in addition to acetylcholine (ACh) and noradrenaline, a transmitter with the same action on thyroid cyclic nucleotide metabolism as VIP.

All experiments were performed on isolated fragments of rat thyroid superfused with modified Krebs–Henseleit solution, gassed with 95% O₂, 5% CO₂ and maintained at 37 °C (ref. 10). Mixed fragments (5–10 mg) of the thyroid glands from 16 rats were placed in 8 small Perspex flow chambers (vol 0.6 ml, flow 1 ml min⁻¹). Four of the chambers contained silver electrodes for electrical field stimulation (FS), the remaining four, without electrodes, served as control ‘partner’ preparations. The preparations (each consisting of the equivalent of two whole thyroid glands) were superfused 30 min before stimulation. At various times following onset of FS or application of drug, both

‡ To whom correspondence should be addressed.

Table 1 Effects of field stimulation (FS) in the presence of TTX on tissue cyclic nucleotide levels

Time	Cyclic AMP test (FS + TTX)	Cyclic AMP control	Cyclic GMP test (FS + TTX)	Cyclic GMP control
10 s	9.38	9.46	2.38	2.46
30 s	6.47	6.50	2.25	2.23
45 s	7.31	7.21	3.36	3.41
1 min	5.97	6.15	1.86	1.79
2 min	7.02	7.20	2.71	2.78
3 min	8.10	8.13	2.58	2.67
5 min	8.05	8.25	3.03	3.00
7 min	6.83	6.86	2.48	2.51
9 min	6.70	6.78	2.12	2.11
12 min	9.15	9.17	3.34	3.41
16 min	7.42	7.35	2.69	2.63
20 min	8.31	8.25	2.15	2.25

Field stimulation was continuous at 20 Hz, 50 V, with pulse durations of 1 ms. Mean ($\pm$ s.e.) control levels of cyclic AMP and cyclic GMP were: 7.61 ± 0.29 and 2.60 ± 0.13 pmol per mg protein, respectively ($n = 12$). The mean ($\pm$ s.e.) levels of cyclic AMP and cyclic GMP following FS in the presence of TTX (10^{-6} M) were 7.56 ± 0.31 and 2.58 ± 0.14 pmol per mg protein, respectively ($n = 12$). The data reveal no statistically significant difference between the test and control values ($P > 0.1$).

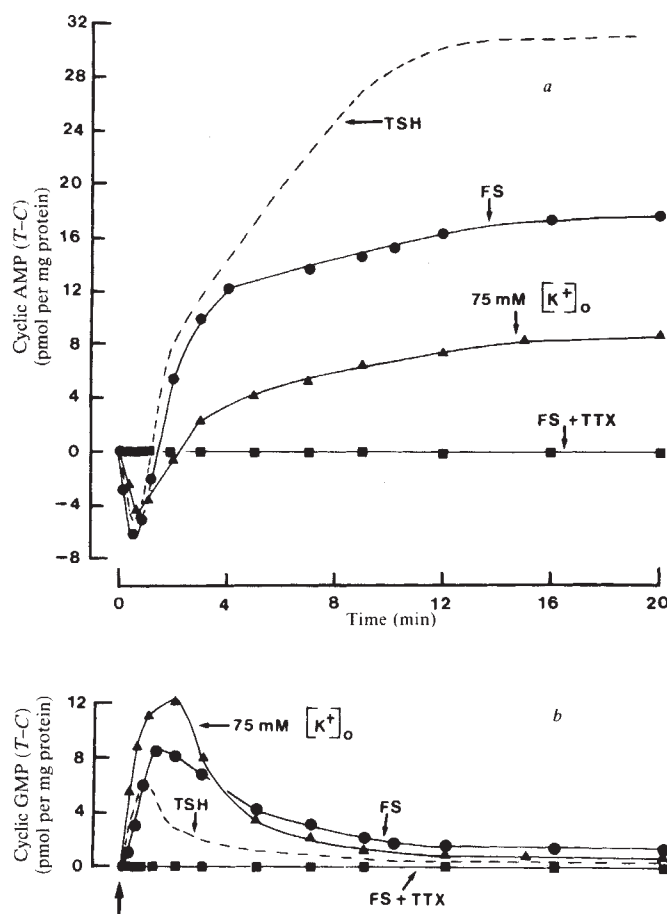


Fig. 1 Time course of changes in the concentration of cyclic AMP (a) and cyclic GMP (b) in superfused rat thyroid segments following electrical field stimulation (FS: 20 Hz frequency, 50 V amplitude, 1 ms pulse width) in the absence (●) and presence (■) of 10^{-6} M TTX and on exposure of preparations to Krebs solution containing 75 mM K^+ (▲). Also shown are the cyclic nucleotide concentrations in the presence of 20 mU ml^{-1} TSH (---). Cyclic AMP and cyclic GMP in control and test frozen tissues were measured at precise time intervals after the start of stimulation using the Radiochemical Centre (Amersham) assay kits TRK 432 and TRK 500, respectively. Total protein was estimated using the Biuret method²¹ and cyclic nucleotide values expressed in pmol per mg protein. Mean ($\pm$ s.e.) control levels of cyclic AMP and cyclic GMP were 7.96 ± 0.18 and 2.38 ± 0.06 pmol per mg protein, respectively ($n = 60$). Each point on the graph represents the difference in cyclic nucleotide concentration between two similar preparations of thyroid segments, one serving as the test and the other as the control.

test and partner preparations (chamber plus tissue) were frozen in liquid nitrogen. Cyclic AMP and cyclic GMP in control and test frozen tissues were extracted and measured¹³.

The cyclic nucleotide concentration in a particular experimental situation was always expressed as the difference between the concentration in the test and the corresponding control preparation. Each data point in Figs 1–3 therefore represents the difference between two single measurements. The reliability of our method can perhaps best be assessed by inspection of Table 1, which shows cyclic nucleotide values in control and test preparations. Because the test in this case consisted of FS plus tetrodotoxin (TTX), that is, nerve stimulation in the presence of a blocker of nerve action potentials, we have, in effect, two parallel controls. Although there is some time fluctuation, it is clear that this occurs in parallel in the test and control preparations.

Figure 1a shows the time-course changes in tissue cyclic AMP levels during FS in the absence and presence of 10^{-6} M TTX and during exposure to 75 mM K^+ . The effect of 20 mU ml^{-1} TSH is also shown for comparison. An initial, rapid, transient fall in cyclic AMP levels occurs in the first 30 s followed by a slower rise reaching almost steady state by 16–20 min. These changes in cyclic AMP concentration were abolished by superfusion of the preparation with TTX. Figure 1b shows time-course changes in tissue cyclic GMP levels during FS, on exposure to 75 mM K^+ and during FS in the presence of 10^{-6} M TTX. The changes evoked by TSH are also shown. The early, rapid fall in cyclic AMP levels is accompanied by a rather slower, transient increase in cyclic GMP levels, maximal after ~1–2 min, then declining slowly, returning to the control level after 10–12 min, and remaining there throughout the stimula-

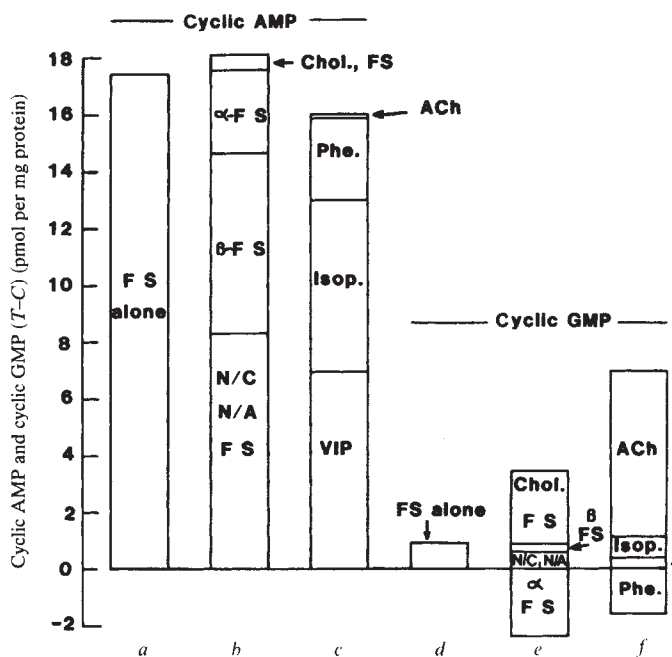


Fig. 2 Histograms showing the effects of electrical field stimulation (FS) and various agonists (for example, Phe., phenylephrine; Isop., isoprenaline; ACh, acetylcholine; and VIP, vasoactive intestinal peptide), in the absence and presence of various combinations of different blockers, on cyclic AMP (a–c) and cyclic GMP (d–f) levels (data from Table 1). a, FS in the absence of blockers. b, N/C, N/A, FS, non-cholinergic and non-adrenergic nervous component (presence of cholinergic and adrenergic blockers); β-FS, β-adrenergic nervous component (presence of cholinergic and α-adrenergic blockers and subtraction of N/C, N/A component); α-FS, α-adrenergic component (presence of cholinergic and β-adrenergic blockers and subtraction of N/C, N/A component). c, VIP (10^{-7} M); Isop. (10^{-5} M); Phe. (10^{-5} M); ACh (10^{-5} M). d, FS in the absence of blockers. e, α-FS, α-adrenergic nervous component; N/C, N/A, non-cholinergic, non-adrenergic nervous component; β-FS, β-adrenergic nervous component; Chol. FS, cholinergic nervous component. f, Phe. (10^{-5} M); VIP (10^{-7} M); Isop. (10^{-5} M); ACh (10^{-5} M).

tion period. TTX (10^{-6} M) abolished the cyclic GMP concentration changes evoked by FS.

The series of experiments summarized in Fig. 1 (thyroids from 200 rats, yielding 50 tests and 50 control preparations) demonstrates directly that electrical stimulation of intrinsic nerves causes marked and complex changes in cyclic nucleotide metabolism. In the steady state (after ~ 15 min stimulation) there is a substantial and sustained elevation in cyclic AMP concentration. As several different transmitters might be involved it was necessary to carry out experiments using various blocking agents. Table 2 shows the results from experiments in which thyroid preparations were superfused for 20 min in the presence or absence of different kinds of stimuli combined with various blockers. The nerve stimulation-evoked increase in tissue cyclic AMP concentration can be explained as the sum of α - and β -adrenergic components in addition to a substantial non-cholinergic, non-adrenergic contribution. The experiments carried out with specific agonist superfusion (isoprenaline, phenylephrine, ACh and VIP) gave results that fit those obtained with the nerve stimulation (Fig. 2). Although there is no unequivocal indication of a cholinergic nervous component in the cyclic AMP results, the results with cyclic GMP make it clear that nerve stimulation releases ACh which acts on muscarinic rather than nicotinic receptor sites (Table 2). In the steady state, nerve stimulation causes only a modest elevation of cyclic GMP concentration because the elevation expected from the cholinergic component is partly balanced by the ability of α -adrenergic stimulation to decrease cyclic GMP levels (Fig. 2).

The finding that the non-cholinergic, non-adrenergic component of the nerve stimulation-evoked increase in cyclic AMP concentration is matched by the effect of VIP (Fig. 2) implicates VIP as the transmitter involved. If this is so, superfusion with VIP should mimic the entire time course of the nerve stimulation (FS)-evoked changes in cyclic AMP and cyclic GMP concentration obtained in the presence of cholinergic and adrenergic blockers. As seen in Fig. 3, this was indeed the case. Figure 3a shows the time-course changes in cyclic AMP levels following FS and exposure to 10^{-7} M VIP and 75 mM K^+ in the presence of atropine, propranolol and phentolamine. There is a large and rapid increase in cyclic AMP, maximal after 30–60 s, followed by an abrupt, transient fall after 2–3 min. Thereafter, the level of cyclic AMP rises gradually to a secondary increase, reaching a steady-state level after 16–20 min. Figure 3b shows time-course changes in cyclic GMP concentration during application of FS, 75 mM K^+ and during exposure to 10^{-7} M VIP. The early rapid rise and fall in cyclic AMP is accompanied by a slow, transient increase in cyclic GMP, reaching a maximum after 2–3 min before declining towards the control level.

The series of experiments summarized in Fig. 3 (thyroids from 136 rats, yielding 34 test and 34 control preparations)

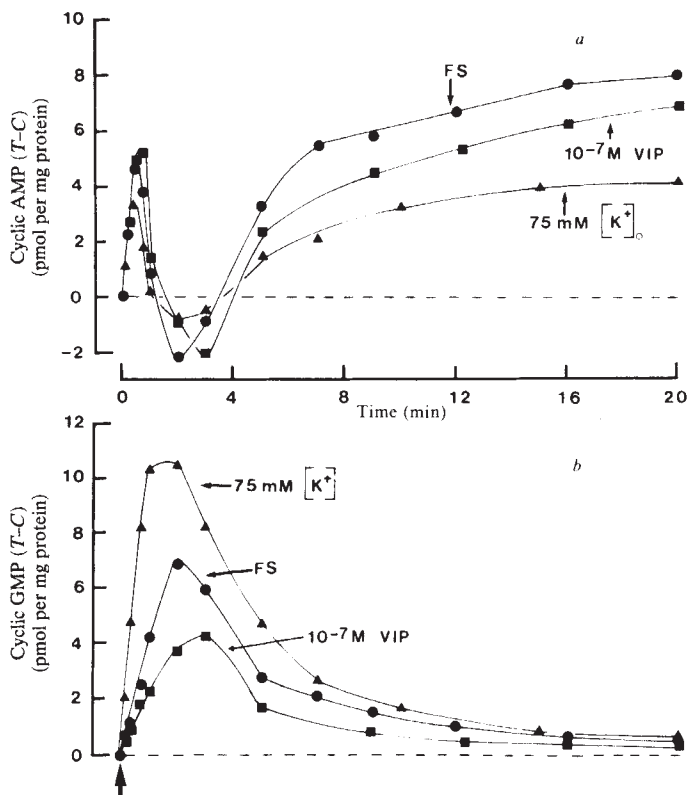


Fig. 3 Time-dependent changes in endogenous cyclic AMP (a) and cyclic GMP (b) levels in rat thyroid segments following FS (●), during application of 75 mM K^+ (▲) and 10^{-7} M VIP (■). Atropine (1.4×10^{-6} M), phentolamine (10^{-5} M) and propranolol (5×10^{-6} M) were present in the superfusing solution from the start of the perfusion period and throughout the period of stimulation. Cyclic AMP and cyclic GMP concentrations are expressed as in Fig. 1.

demonstrates that the non-cholinergic, non-adrenergic component of the nerve-stimulation response has the opposite initial effect on cyclic AMP levels to that evoked by nerve stimulation in the absence of blockers (Fig. 1). The complex triphasic cyclic AMP response to non-cholinergic, non-adrenergic nerve stimulation is remarkably well reproduced by VIP stimulation. The response is clearly not a non-specific peptide effect because at steady state (20 min), the structurally very different peptide caerulein, used in the same concentration as VIP (10^{-7} M), decreases the tissue cyclic AMP concentration (Table 2).

Elevation of extracellular K^+ concentration is generally accepted as a powerful stimulant of neurotransmitter and hormone secretion in a variety of tissues^{13–16}. Figures 1 and 3 demonstrate that high K^+ stimulation does to a large extent mimic the effect of nerve stimulation, although the cyclic GMP

Table 2 Effects of different stimuli (20 min) in the presence or absence of various blocking agents on thyroid cyclic nucleotide levels

	Cyclic AMP			Cyclic GMP		
	Test (T)	Control (C)	(pmol per mg protein) T - C	Test	Control	T - C
a FS alone	25.74	8.25	17.49	3.11	2.25	0.86
b FS + atropine, propranolol, phentolamine	16.45	8.20	8.25	2.84	2.25	0.59
c FS + atropine + phentolamine	20.77	6.15	14.62	3.17	2.36	0.81
d FS + phentolamine + propranolol	14.88	6.15	8.73	5.52	2.36	3.16
e FS + atropine + propranolol	17.36	6.15	11.21	0.55	2.36	-1.81
f FS + atropine + phentolamine + propranolol + hexamethonium	16.05	7.93	8.12	2.38	2.05	0.33
g ACh (10^{-5} M) + phentolamine + propranolol	7.03	6.97	0.06	8.09	2.26	5.83
h ACh (10^{-5} M) + hexamethonium	6.92	6.98	-0.06	8.17	2.97	5.20
i ACh (10^{-5} M) + atropine	6.12	6.05	0.07	2.85	2.65	0.20
j Phenylephrine (10^{-5} M) + atropine + propranolol	9.62	6.79	2.83	0.79	2.38	-1.59
k Isoprenaline (10^{-5} M) + atropine + phentolamine	12.98	6.97	6.01	2.94	2.26	0.68
l VIP (10^{-7} M) + atropine + phentolamine + propranolol	14.82	7.90	6.92	2.67	2.35	0.32
m Caerulein (10^{-7} M) + propranolol + atropine + phentolamine	5.21	9.67	-4.46	3.99	2.47	1.52
n TSH (20 mU ml ⁻¹)	37.24	6.28	30.96	3.00	2.61	0.39

The concentrations of blockers used were (M): atropine, 10^{-6} ; propranolol, 5×10^{-6} ; phentolamine, 10^{-5} ; hexamethonium, 10^{-5} . Each value listed under control and test is the result of a single measurement.

changes are more marked than with electrical stimulation and the effects on cyclic AMP metabolism are weaker. High K^+ will, in addition to depolarizing intra-thyroidal nerve endings, depolarize the follicular cell membrane^{17,18}, but this is not accompanied by any change in the membrane resistance¹⁸, as the thyroid follicular cell membrane is electrically inexcitable¹⁷. All effects of FS are abolished by TTX¹⁹. As thyroid follicular cells do not fire action potentials^{17,18}, the only specific site for TTX action is on intra-thyroidal nerves. Furthermore, ACh, adrenaline, VIP and TSH have no effect on thyroid follicular cell membrane potential and resistance^{17,18}.

The present work has shown that two recognized methods of stimulating nerves *in vitro* elicit similar changes in thyroid cyclic nucleotide levels, lending powerful support to the hypothesis that intra-thyroidal nerves can influence thyroid follicular cells directly. The pattern of the innervation is complex. Clearly, further work is required before the neurotransmitters responsible can be definitely identified, but the results presented here do show functional cholinergic, adrenergic as

well as non-cholinergic, non-adrenergic innervation of the thyroid gland. As exogenous VIP closely mimics the non-cholinergic, non-adrenergic nerve response (Fig. 3) and VIP-immunoreactive nerves have been demonstrated in the thyroid⁹, it seems likely that the neurotransmitter is VIP or a closely related peptide. Nerve-mediated mechanisms in the control of thyroid function could constitute a sensitive and rapidly responding mechanism, allowing accurate modification of glandular metabolism to accommodate fluctuating requirements of the organism.

The finding that all the stimulation procedures resulted in changes in both cyclic AMP and cyclic GMP levels emphasizes the danger of focusing exclusively on cyclic AMP metabolism. The cyclic nucleotide changes generally display a reciprocal pattern which has also been observed in the frog heart²⁰ and the pancreas¹³. The functional implication of this complex inter-relationship obviously deserves much further attention.

This work was supported in part by the MRC. We thank Mr B. Weryk and Mrs K. Kleppang for technical assistance.

Received 16 July 1981; accepted 3 March 1982.

- Dumont, J. E. *Vitam. Horm.* **29**, 287–412 (1971).
- Tong, W. in *Handbook of Physiology* Vol. 3, Sect. 7 (eds Greer, M. A. & Solomon, D. H.) 255–283 (American Physiological Society, Washington DC, 1974).
- Field, J. B. in *The Thyroid* 4th edn (eds Werner, S. C. & Ingbar, S. H.) 185–195 (Harper & Row, New York, 1978).
- Söderberg, U. *Acta physiol. scand.* **42**, Suppl. 147, 1–113 (1958).
- Melander, A. in *The Thyroid* 4th edn (eds Werner, S. C. & Ingbar, S. H.) 216–222 (Harper & Row, New York, 1978).
- Dumont, J. E. *et al. Adv. Cyclic Nucleotide Res.* **9**, 723–734 (1978).
- Melander, A., Sundler, F. & Westgren, U. *Endocrinology* **96**, 102–106 (1975).
- Melander, A. & Sundler, F. *Endocrinology* **105**, 7–9 (1979).
- Ahrens, B. *et al. Nature* **287**, 343–345 (1980).
- Davison, J. S., Pearson, G. T. & Petersen, O. H. *J. Physiol., Lond.* **301**, 295–305 (1980).
- Nishiyama, A., Kato, K., Saitoh, S. & Wakui, M. *Membrane Biochem.* **3**, 49–66 (1980).
- Pearson, G. T., Davison, J. S., Collins, R. C. & Petersen, O. H. *Nature* **290**, 259–261 (1981).
- Pearson, G. T., Singh, J., Daoud, M. S., Davison, J. S. & Petersen, O. H. *J. biol. Chem.* **256**, 11025–11031 (1981).
- Fenn, W. O. *Physiol. Rev.* **20**, 377–415 (1940).
- Schramm, M. *Biochim. biophys. Acta* **165**, 546–549 (1968).
- Douglas, W. W. *Br. J. Pharmac.* **34**, 451–474 (1968).
- Green, S. T. & Petersen, O. H. *Pflügers Arch. ges. Physiol.* **391**, 119–124 (1981).
- Green, S. T., Singh, J. & Petersen, O. H. *Biochim. biophys. Acta* **720**, 36–41 (1982).
- Narahashi, T. *Physiol. Rev.* **54**, 813–889 (1974).
- Flitney, F. W. & Singh, J. *J. Physiol., Lond.* **304**, 21–42 (1980).
- Gornall, A. G., Bardawill, C. J. & David, M. B. *J. biol. Chem.* **177**, 751–766 (1949).

A differentiation factor required for the expression of cytotoxic T-cell function

David H. Raulet* & Michael J. Bevan*

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Effector cytotoxic T lymphocytes (CTLs) are induced from resting precursors (small lymphocytes) which do not express cytotoxic activity. The process of CTL activation, which occurs in 2–7 days following mitogen or antigen stimulation, thus involves a differentiation event and normally also involves proliferation. These events in CTL activation seem to be induced, at least in part, by soluble, antigen-nonspecific mediators which reach the CTL precursor as a consequence of cell–cell interactions during the immune response. The initial evidence for this was the finding that in addition to CTL precursors, Ia-positive macrophage-like accessory cells^{1–4}, and helper T cells^{5–10} are required for the induction of CTL activity *in vitro*. Furthermore, the requirement for accessory cells can be replaced with factors present in concanavalin A (Con A)-induced cultures of spleen cells¹¹. There is evidence that T-cell products such as interleukin-2 (IL-2, formerly T-cell growth factor)^{11–15} as well as accessory cell products, such as interleukin-1 (IL-1, formerly lymphocyte activating factor)^{4,15}, are involved in T-cell activation, although the precise roles of these factors in CTL activation are poorly understood. Here we provide evidence for a factor, different from IL-1 and IL-2, which is present in supernatants of Con A-activated mouse spleen cells and is required for the differentiation of CTL precursors into active CTLs.

The experimental system used involved the Con A-mediated polyclonal induction of CTLs from T-cell populations¹⁶. The responding T-cell population was depleted of accessory cells so that Con A-induced proliferation¹⁷ and differentiation were

dependent on exogenous factors. CTLs were assayed after 40 h for cytotoxic activity against ⁵¹Cr-labelled P815 targets in the presence of phytohaemagglutinin (PHA); this assay registers the activity of all CTLs, regardless of their antigen specificity^{16,18}. We wished to determine whether IL-2, which is assayed by its capacity to maintain the proliferation of cloned CTLs¹⁹, is sufficient to support the Con A-induced differentiation of CTLs from accessory cell-depleted T-cell populations. In initial experiments we compared two sources of IL-2: supernatants from Con A-activated spleen cell cultures (spleen Con A Sn)^{12,13} and supernatants from cultures of Con A-activated T-cell hybridoma AOFS 21.10.9 cells (AOFS Con A Sn). AOFS 21.10.9 (provided by Drs J. Kappler and P. Marrack) is a cloned, murine T-cell hybridoma^{20,21} which releases IL-2 on stimulation with Con A or antigen.

Our initial observation was that T cells activated with Con A plus AOFS Con A Sn were induced to proliferate (Fig. 1b), but did not mature into active CTLs after 40 h (Fig. 1a). In contrast, T cells activated with Con A plus spleen Con A Sn were efficiently induced to differentiate into cytotoxic cells

Table 1 Lyt-2-positive T cells proliferate in response to Con A plus AOFS Con A Sn

Additions to test culture	³ H-thymidine (c.p.m. × 10 ⁻² ± s.e.) per culture incorporated by:	
	T cells	Lyt-2-positive T cells
—	0.3 ± 0.1	0.2 ± 0.1
Con A (2 µg ml ⁻¹)	3.4 ± 0.6	1.6 ± 0.2
Con A + 25% spleen Con A Sn	152.4 ± 4.1	357.9 ± 10.4
Con A + 12% spleen Con A Sn	108.2 ± 3.2	246.5 ± 11.6
Con A + 25% AOFS Con A Sn	159.1 ± 9.2	336.8 ± 25.6
Con A + 12% AOFS Con A Sn	99.4 ± 3.2	307.8 ± 27.5

T cells were purified as described in Fig. 1 legend. Lyt-2-positive T cells were purified from accessory cell-depleted T-cell populations by positive selection on plastic dishes³⁰ coated with monoclonal anti-Lyt-2 antibodies (N. Landau and M.J.B., in preparation). 60 × 10⁶ T cells in 4 ml were incubated for 40 min in the cold on a 100-mm Petri dish which had been coated for 1 h in the cold with 20 µg of purified AD4(15)³¹ anti-Lyt-2 antibodies. The non-adherent cells were removed, the plates were washed six times and the adherent cells (17% of applied cells in this experiment) were removed from the dish by pipetting. T-cell populations selected in this fashion are ≥95% Lyt-2⁺. The proliferative response of T cells and Lyt-2⁺ T cells were determined as in Fig. 1 legend.

* Present addresses: Department of Pathology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA (D.H.R.); Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, California 92037, USA (M.J.B.).

Fig. 1 IL-2 is not sufficient to support the differentiation of CTLs from accessory cell-depleted T-cell populations. Spleen Con A Sn^{12,13} was prepared by culturing spleen cells (7×10^6 cells ml^{-1}) from BALB.K mice with $2 \mu\text{g ml}^{-1}$ Con A for 24 h in RPMI-1640 medium containing 5% heat-inactivated fetal calf serum, $50 \mu\text{M}$ 2-mercaptoethanol, 0.02% glutamine, 5 mM HEPES, $50 \mu\text{g ml}^{-1}$ gentamycin, 100 U ml^{-1} penicillin and $100 \mu\text{g ml}^{-1}$ streptomycin (complete medium). AOFS Con A Sn^{20,21} was prepared by activating AOFS-21.10.9 cells (5×10^5 cells ml^{-1}) with $10 \mu\text{g ml}^{-1}$ Con A for 24 h in complete medium. The cells were removed by centrifugation and the supernatants sterilized by filtration. Excess Con A was removed from AOFS Con A Sn by passage through a column of Sepharose-4B which had been coupled with *p*-aminophenyl- α -D-mannopyranoside. Concentrated AOFS Con A Sn was prepared from AOFS Con A Sn (prepared as above,

except with 0.5% fetal calf serum and $5 \mu\text{g ml}^{-1}$ Con A) by precipitation with ammonium sulphate. The fraction that precipitated between 40 and 80% ammonium sulphate, which contains virtually all the IL-2 activity, was resuspended in $1/30$ the original volume, and dialysed extensively against RPMI-1640 medium; the concentrations of concentrated AOFS depicted as based on the original volume. Accessory cell-depleted T-cell populations were prepared by passing 5×10^8 lymph node cells from BALB.K mice through a column containing 3.5 g of washed nylon wool³², and treating the non-adherent cells with ATH anti-ATL serum (anti-Ia^k serum) and rabbit complement essentially as described previously³¹. Surviving cells ('T cells') were washed twice before use. The capacity of the supernatants to induce the differentiation of CTLs from accessory cell-depleted T-cell populations was determined by incubating aliquots of 4×10^6 T cells in wells of Costar 3524 plates with $2 \mu\text{g ml}^{-1}$ Con A and the indicated concentration of supernatant in a total volume of 1.5 ml complete medium. After 40 h at 37°C in an atmosphere of $5\% \text{ CO}_2$ in air, the lectin was dissociated from the activated cells by incubating them with α -methyl mannoside (50 mM) for 10 min at room temperature. The cells were washed once, resuspended in 1 ml and serial dilutions of this assayed for cytotoxicity against ^{51}Cr -labelled P815 targets in the presence of $10 \mu\text{g ml}^{-1}$ PHA-P¹⁶ (Difco) for 6 h essentially as described previously³¹. Lytic units are a relative measure of CTL activity, and are calculated from complete titrations of effector cells versus the per cent specific lysis of target cells. One lytic unit³³ is the number of CTLs which will cause 30% specific lysis of 4×10^4 target cells in this assay period. Comparable numbers of cells were recovered from Con A-activated cultures supplemented with either spleen Con A Sn or AOFS Con A Sn. Con A-dependent proliferation of accessory cell-depleted T-cell populations was determined by incubating 2×10^5 cells in 0.2 ml in Costar 3596 microwells with $2 \mu\text{g ml}^{-1}$ Con A and the indicated concentration of supernatant for 40 h at 37°C , at which time the cultures were pulsed for 5 h with $2 \mu\text{Ci } ^3\text{H}$ -thymidine³¹. Results shown are the mean of triplicate cultures $\pm \text{s.e.}$ IL-2 activity was determined essentially as described elsewhere¹⁹. Cultures (0.2 ml) containing 1×10^4 clone-3 cells, serial dilutions of supernatants and 25 mM α -methyl mannoside to block residual Con A, were incubated at 37°C for 20 h, and the cultures then pulsed with ^3H -thymidine. Clone-3 (provided by Dr Thomas Hünig) is a cloned, IL-2-dependent CTL line from a CBA mouse. Supernatants tested were spleen Con A Sn ($\circ$), AOFS Con A Sn ($\blacksquare$), AOFS Con A Sn concentrated by ammonium sulphate precipitation (Δ), a mixture of 70% concentrated AOFS Con A Sn and 70% spleen Con A Sn (∇), and no supernatant ($\otimes$).

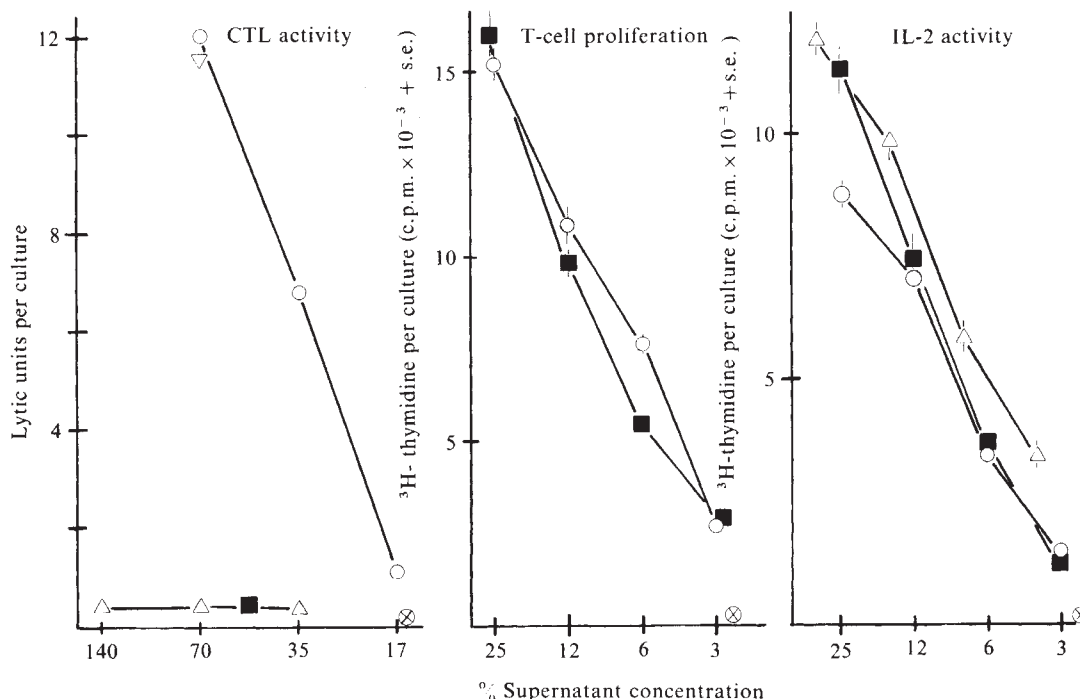


Fig. 1a, b). (Cytotoxic cells induced with Con A and spleen Con A Sn were sensitive to complement-mediated lysis with monoclonal anti-Thy-1 or anti-Lhy-2 antibodies, and are thus conventional CTLs (data not shown).) As can be determined from the titration curves shown in Fig. 1c, the two kinds of supernatant had comparable levels of IL-2, and of factors which induce the proliferation of Con A-activated accessory cell-depleted T cells (Fig. 1b). In addition, AOFS Con A Sn did not suppress the induction of CTLs by spleen Con A Sn when included in the same culture (Fig. 1a). These results suggest that a factor(s) other than IL-2, present in spleen Con A Sn but not in AOFS Con A Sn, is required for CTL differentiation. We call this factor 'CTL differentiation factor' (CDF).

Because AOFS Con A Sn induces T-cell proliferation without inducing differentiation into CTLs, it seemed possible that CTL proliferation and differentiation were controlled by separate factors, IL-2 and CDF, respectively. However, although AOFS Con A Sn supports the growth of CTL lines, it was possible that CTL precursors do not proliferate in response to Con A plus AOFS Con A Sn.

To ascertain whether AOFS Con A Sn plus Con A induces the proliferation of Lyt-2-positive T cells (that is, the T-cell subset which contains CTL precursors²²), we purified Lyt-2-positive T cells by absorption to plastic dishes coated with monoclonal anti-Lyt-2 antibodies, and compared the proliferative responses of these cells with the responses of unselected T cells. The proliferative response of Lyt-2-positive T cells to Con A plus AOFS Con A Sn was actually larger than the com-

parable response of unselected T cells (Table 1). In a related experiment, we determined whether Con A plus AOFS Con A Sn induced the blast transformation of Lyt-2-positive T cells in cultures of unseparated T cells. Density step gradients were used to isolate blast cells (activated cells) from 40-h cultures of accessory cell-depleted T-cell populations activated with Con A plus AOFS Con A Sn. The proportion of these T cells susceptible to lysis with monoclonal anti-Lyt-2.2 antibodies plus complement (27%) did not differ significantly from the proportion of Lyt-2-positive blast cells in cultures activated with Con A plus spleen Con A Sn (32%). Thus, we find no evidence that AOFS Con A Sn is less effective in inducing the proliferation or blast transformation of CTL precursors than is spleen Con A Sn.

In further experiments we determined some of the physico-chemical properties of CDF. Spleen Con A Sn was treated in various ways, and then tested for CDF and IL-2 activities. Dialysis against culture medium resulted in a minimal loss of CDF activity, which shows that CDF is a macromolecule (Table 2). Furthermore, treatment of CDF-containing supernatants with trypsin abolished CDF activity, even when the digested sample was supplemented with AOFS Con A Sn for the CDF assay. Thus, CDF has a polypeptide component which is essential for activity. Dialysis of spleen Con A Sn against an acidic ($\text{pH } 2.0$) buffer abolished CDF activity but had no effect on IL-2 activity. Thus, unlike IL-2²³, CDF is acid sensitive.

The finding that CDF activity is abolished by digestion with trypsin is one piece of evidence that CDF is not identical to

Table 2 Physicochemical properties of CDF

Expt	Treatment of Sn	CDF activity % Specific lysis by CTLs induced with indicated concentration of test Sn			IL-2 activity ³ H-thymidine (c.p.m. $\pm$ s.e.) incorporated in clone-3 cells with the indicated concentration of test Sn		
		50%	25%	12%	25%	12%	6%
1	Untreated	32	19	6	8,004 $\pm$ 58	6,244 $\pm$ 261	3,715 $\pm$ 203
	Dialysis vs. medium	21	12	4	10,802 $\pm$ 397	7,091 $\pm$ 400	3,805 $\pm$ 312
	Dialysis vs. pH 2	2	2	—	8,598 $\pm$ 96	7,257 $\pm$ 44	3,509 $\pm$ 750
2	Untreated	46	27	8	8,929 $\pm$ 299	8,016 $\pm$ 234	7,118 $\pm$ 795
	Dialysis vs. medium	22	18	6	8,218 $\pm$ 452	7,205 $\pm$ 181	6,674 $\pm$ 422
	Untreated	29	25	—	11,743 $\pm$ 795	10,278 $\pm$ 719	5,719 $\pm$ 313
3	Trypsin treated	3	2	—	297 $\pm$ 83	469 $\pm$ 81	—
	Trypsin treated, restored with AOFs Sn	6	2	—	—	—	—
	Control for trypsin treatment	39	23	6	10,319 $\pm$ 116	9,123 $\pm$ 722	5,141 $\pm$ 478
	AOFs Sn	3	—	—	—	—	—
	—	—	—	—	—	—	—

For expts 1 and 2, 10-ml samples of spleen Con A Sn were either dialysed against two changes (250 ml each) of complete medium (see Fig. 1 legend) or dialysed against 0.1 M glycine, 0.15 M NaCl, pH 2.0, for 18 h, followed by dialysis against two changes (250 ml each) of complete medium. For expt 3, we used serum-free spleen Con A Sn prepared as described in Fig. 1 legend, except that the medium was supplemented with $10 \mu\text{g ml}^{-1}$ of bovine serum albumin instead of serum. Samples of serum-free spleen Con A Sn were treated with $20 \mu\text{g ml}^{-1}$ trypsin (TPCK treated, Millipore) for 1.5 h at 37°C , at which time fetal calf serum (which contains trypsin inhibitors) was added to a final concentration of 5%. As a control (expt 3, 4th line) a sample of serum-free spleen Con A Sn was incubated in identical conditions, except that trypsin was added only after the addition of serum. In a separate experiment, treatment of supernatants with Difco trypsin also abolished CDF and IL-2 activities (not shown). Treated supernatants were sterilized by filtration and tested for CDF and IL-2 activities as described in Fig. 1 legend, except that cytotoxic activity is depicted in terms of the % specific lysis of target cells at a single effector to target ratio (25:1). In one case (expt 3, 3rd line), the test supernatant was supplemented with concentrated AOFs Con A Sn (to an adjusted concentration of 140%) for the CDF assay.

IL-1, which is trypsin insensitive²⁴. In addition, we have tested IL-1 preparations from two sources (Dr S. Mizel, and Drs J. Kappler and P. Marrack) in our assay: both preparations were negative for CDF activity, even if supplemented with excess AOFs Con A Sn before the test (data not shown). Taken together, our results suggest that CDF activity resides in a molecular species that is distinct from IL-1 and IL-2.

To demonstrate further the molecular independence of IL-2 and CDF, we attempted to remove IL-2 selectively from spleen Con A Sn in such a way that the supernatant retained CDF activity. IL-2 was depleted from spleen Con A Sn by absorption with cells of an IL-2-dependent cloned CTL line, CTLL-15H (ref. 25). Following absorption, the supernatants were dialysed and tested for CDF and IL-2 activities. Although virtually all IL-2 was removed (Fig. 2b), the absorbed supernatants induced the differentiation of CTLs with ~50% the efficiency of unabsorbed spleen Con A Sn (Fig. 2a). A mixture of the absorbed spleen Con A Sn and AOFs Con A Sn was as effective as unabsorbed spleen Con A Sn in inducing CTL activity (Fig. 2a). From these data, we conclude that CDF is separable from IL-2;

furthermore, because the absorbed supernatant had no measurable IL-2 activity and did not inhibit the IL-2 assay significantly (Fig. 2b), it is unlikely that CDF is simply a modified form of IL-2 with both IL-2 and CDF activities.

It is interesting that we find no requirement for exogenous IL-2 in this assay system (Fig. 2a), particularly as IL-2 is implicated in most models of CTL differentiation^{6,7,11-14}. In fact, our data do not rule out the involvement of IL-2 in CTL differentiation, because it is possible that Con A-activated T cells in the test culture produce IL-2 in response to factors in the IL-2-depleted spleen Con A Sn.

The absence of a requirement for factors other than IL-2 in CTL induction in other published reports may be because in such studies, spleen Con A Sn was the sole source of IL-2, and biochemical separation of IL-2 from other factors in spleen Con A Sn is often difficult. In addition, we have found that with extended culture periods (4–5 days) CTL activity appears in cultures to which no exogenous CDF was added (data not shown). This may be due to *in situ* production of CDF by cells in the test culture.

We do not know whether CDF is required to maintain the CTL activity of mature CTLs. The factor does not seem to bind avidly to cloned, mature CTLs, because in conditions in which IL-2 is depleted from spleen Con A Sn by absorption to IL-2-dependent, cloned CTLs, CDF activity is not depleted (Fig. 2).

Recent studies of soluble factors involved in the antibody response have provided evidence for separate factors which regulate the growth and differentiation to immunoglobulin secretion of anti-immunoglobulin or antigen-activated B cells²⁶⁻²⁹. Our data suggest that CTL activation may also involve separate growth-inducing and differentiation-inducing activities—IL-2 and CDF, respectively. Further work will be required to determine the cellular source and specific mechanism of action of CDF.

This work was supported by USPHS grant AI-14269. We thank Lorraine Gemmell for technical assistance, Fern Brown for typing the manuscript, and Dr David Parker for performing the absorption of Sp Con A Sn with CTLL cells.

Received 24 September 1981; accepted 22 February 1982.

- Wagner, H., Feldman, M., Boyle, W. & Schrader, J. W. *J. exp. Med.* **136**, 331–340 (1972).
- Koren, H. S. & Hodes, R. J. *Eur. J. Immun.* **7**, 394 (1977).
- Taniyama, T. & Holden, H. T. *J. Immun.* **123**, 43–49 (1979).
- Farrar, W. L., Mizel, S. B. & Farrar, J. J. *J. Immun.* **124**, 1371–1377 (1980).
- Cantor, H. & Boyse, E. A. *J. exp. Med.* **141**, 1390–1399 (1975).
- Von Boehmer, H. & Haas, W. *Immun. Rev.* **54**, 27–56 (1981).
- Wagner, H. *et al. Immun. Rev.* **51**, 215 (1980).
- Cooley, M. A. & Schmitt-Verhulst, A.-M. *J. Immun.* **123**, 2328–2336 (1979).
- Glasebrook, A. & Fitch, F. J. *exp. Med.* **151**, 876–895 (1980).
- Finberg, R., Green, M., Benacerraf, B. & Burakoff, S. J. *J. Immun.* **123**, 1205–1209 (1979).
- Farrar, W. L., Johnson, H. M. & Farrar, J. J. *J. Immun.* **126**, 1120–1125 (1981).

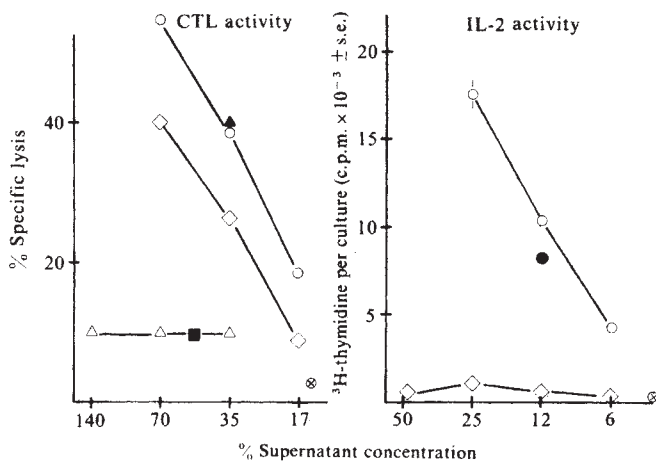


Fig. 2 Spleen Con A Sn, depleted of IL-2, induces CTL differentiation. IL-2 was depleted from spleen Con A Sn by absorption with 2×10^6 CTLL-15H²⁵ (IL-2-dependent CTL line) cells for 24 h at 37°C . Cells were removed by centrifugation and the supernatant was dialysed against two changes of 250 ml complete medium and sterilized by filtration. T cells were prepared and assays were performed as for Fig. 1, except that instead of lytic units CTL activity is expressed as the % specific lysis at an effector to target ratio of 25:1. Supernatants tested were spleen Con A Sn (○), IL-2-depleted spleen Con A Sn (◇), AOFs Con A Sn (■), AOFs Con A Sn concentrated by ammonium sulphate precipitation (% Sn based on original volume) (△), a mixture of 35% IL-2-depleted Con A Sn and 30% concentrated AOFs Sn (▲), a mixture of 12% spleen Con A Sn and 25% IL-2 depleted spleen Con A Sn (●), and no supernatant (⊙).

12. Morgan, D. A., Ruscetti, F. W. & Gallo, R. C. *Science* **193**, 1007–1008 (1976).
13. Gillis, S. & Smith, K. A. *Nature* **268**, 154–156 (1977).
14. Schreier, M. H., Iscove, N. N., Tees, R., Aarden, L. & Von Boehmer, H. *J. Immun. Rev.* **51**, 315–336 (1980).
15. Larsson, E.-L., Iscove, N. N. & Coutinho, A. *Nature* **283**, 664–666 (1980).
16. Bevan, M. J. & Cohn, M. *J. Immun.* **114**, 559–565 (1975).
17. Habu, S. & Raff, M. C. *Eur. J. Immun.* **7**, 451–457 (1977).
18. Möller, G., Sjöberg, O. & Andersson, J. *Eur. J. Immun.* **2**, 586 (1972).
19. Gillis, S. & Smith, K. A. *J. Immun.* **120**, 2027–2032 (1978).
20. Harwell, L., Skidmore, B., Marrack, P. & Kappler, J. *J. exp. Med.* **152**, 893–904 (1980).
21. Kappler, J. W., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198–1214 (1981).
22. Cantor, H. & Boyse, E. A. *J. exp. Med.* **141**, 1376–1389 (1975).
23. Mochizuki, D., Watson, J. & Gillis, S. *J. Immun.* **125**, 2579–2583 (1980).
24. Mizel, S. B. *J. Immun.* **122**, 2167–2172 (1979).
25. Baker, P. E., Gillis, S. & Smith, K. A. *J. exp. Med.* **149**, 273 (1979).
26. Parker, D. C. *Immun. Rev.* **52**, 115 (1980).
27. Swain, S. L., Dennert, G., Warner, J. F. & Dutton, R. W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2517–2521 (1981).
28. Parker, D. C. *J. Immun.* (submitted).
29. Liebson, J., Marrack, P. & Kappler, J. *J. exp. Med.* **154**, 1681 (1981).
30. Wysocki, L. J. & Sato, V. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2844 (1978).
31. Raulet, D. H., Gottlieb, P. D. & Bevan, M. J. *J. Immun.* **125**, 1136–1143 (1980).
32. Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3**, 645–649 (1973).
33. Cerottini, J. C., Engers, H. D., MacDonald, H. R. & Brunner, K. T. *J. exp. Med.* **140**, 703 (1974).

brilliant blue, revealed only the three interferon bands, corresponding to molecular weights of 35,000, 28,000 and 22,000, and no contaminating bands. The 35,000- and 28,000-molecular weight species correspond to mouse IFN- β and the 22,000 species to mouse IFN- α (ref. 8). The purified interferon was stabilized with 0.5% BALB/c serum.

Intravenous injection of 1,000 U (Fig. 1c) or 100 U (Fig. 1d) interferon 2 h after the first sensitization with an antigen overload (50 μ l DNFB) increased ear swelling compared with untreated or buffer-treated controls. Ear swelling was not diminished in animals treated with interferon 2 h after optimal sensitization (Fig. 1f–j), which shows that T_{DH} cells were not affected by the amount of interferon used. The enhancement of T_S -mediated suppression of contact sensitivity by interferon after antigen overload suggests that interferon inhibits the manifestation of the T_S response as discussed previously¹.

In the following experiments (Figs 2–4) T_S were induced by i.v. injection of 15 mg DNBSO₃ per mouse^{2,5,9}. To confirm the presence of T_S , spleen cells (1×10^8 cells per mouse) prepared 6 days after DNBSO₃ injection from donor spleens were transferred to naive recipients. They were sensitized on two consecutive days with an optimal amount of DNFB and challenged 5 days later. Intravenous injection of 1,000 U interferon 2 h after i.v. injection of DNBSO₃ inhibited the generation of T_S (Fig. 2d); thus no suppression could be transferred with the spleen cells of interferon-treated animals to the recipients as compared with the untreated (Fig. 2b) or buffer-treated (Fig. 2c) controls. One hundred units of interferon (Fig. 2e) were marginally effective, ten units (Fig. 2f) had no effect. In the following experiment the effect of interferon on T_S in the recipient animals was investigated by i.v. injection of the recipients with interferon 2 h after T_S transfer (Fig. 3). The activity of the transferred T_S was completely abolished by treating the recipient animals with 1,000 U interferon (Fig. 3d). Inhibition of T_S activity has also been observed in recipient animals injected with *C. parvum* or serum obtained from *C. parvum*-treated animals^{2,3}.

Interferon inhibits the suppressor T cell response of delayed-type hypersensitivity

J. Knop*, R. Stremmer*, C. Neumann*,
E. De Maeyer† & E. Macher*

* Hautklinik der Universität Münster, von-Esmarch-Strasse 56,
D-4400 Münster, FRG

† Institut Curie, Section de Biologie, Centre Universitaire,
F-91405 Orsay, France

We have recently shown that pretreatment of BALB/c mice with *Corynebacterium parvum* preferentially inhibits suppressor T lymphocytes (T_S) which have been induced by an epicutaneous antigen overload¹ or by intravenous (i.v.) injection of 2,4-dinitrobenzenesulphonic acid (DNBSO₃)². Serum obtained 24 h after *C. parvum* injection had a similar suppressive effect on T_S but no effect on T effector lymphocytes of delayed-type hypersensitivity (T_{DH})³. Inhibition of T_S by this serum could be neutralized by an anti-interferon globulin⁴. A crude mouse fibroblast interferon (IFN) preparation also inhibited the T_S response⁴, which suggested that IFN in *C. parvum* serum might be the T_S inhibitory factor. We have now studied the effect of an electrophoretically purified mouse interferon (IFN- α and β) preparation on the T_S response to dinitrofluorobenzene (DNFB) in BALB/c mice. Small amounts (1,000 U) of the preparation inhibited the generation of T_S by DNBSO₃ *in vivo*, and the function of T_S both in the recipient animals when injected i.v. 2 h after transfer of the T_S and after incubation of the spleen cells from tolerant animals with interferon before transfer into recipients. The T_{DH} response was not affected by this amount of interferon.

BALB/c mice were sensitized to DNFB by the method of Phanuphak *et al.*^{1,5}. Briefly, 15 μ l of 0.5% DNFB were painted on to the shaved abdominal skin of BALB/c mice on two consecutive days; 5 days later one ear was challenged with 0.3% DNFB. Ear swelling was measured 24 h later using an engineer's micrometer and expressed as 10^{-2} mm difference in thickness between the unchallenged and challenged ears. Increasing the sensitizing dose of DNFB (50 μ l 0.5%) resulted in diminished contact sensitization¹. Sy *et al.*⁶ have shown that the diminished immune response is due to the generation of T_S by the antigen overload.

Interferon was prepared in mouse C-243 cells by induction with Newcastle disease virus. It was purified by sequential chromatography on Sepharose-bound poly(U) and anti-interferon globulin as previously described⁷. Electrophoretic analysis on polyacrylamide gel, followed by staining with Coomassie

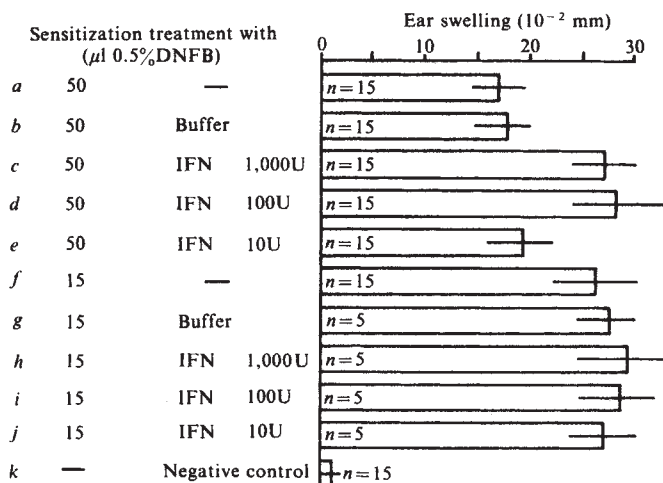


Fig. 1 Effect of interferon on contact sensitivity in animals sensitized with an optimal amount of DNFB (15 μ l) and by epicutaneous antigen overload (50 μ l). BALB/c mice, 10–12 weeks old, were sensitized with 15 μ l (optimal) or 50 μ l 0.5% DNFB by painting the shaved abdominal skin on two consecutive days. Interferon or buffer (0.5 ml per mouse) was injected i.v. once 2 h after the first sensitization. On day 5, ear challenge was performed by painting the right ear with 20 μ l 0.3% DNFB; 24 h later ear thickness was measured and ear swelling was calculated as the difference in ear thickness between the unchallenged (left) and challenged (right) ear. The results of three experiments are presented in this figure and in Figs 2–4 as mean ear swelling $\pm$ s.d. All data were analysed by one-way analysis of variance followed by Scheffe's test for least significant difference. Statistical analysis: group a not different from b; group c increased over a and b; group d increased over a and b; group e not different from a and b; group f not different from g, h, k, j; group k (nonsensitized animals, ear challenged) less than all other groups.

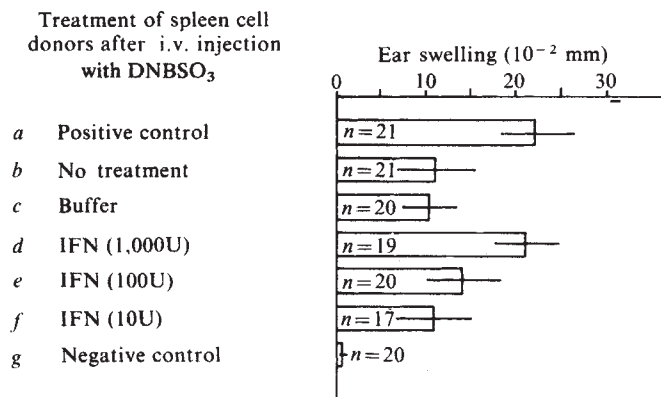


Fig. 2 Effect of interferon on the T_S response in the spleen cell donor treated with DNBSO₃ i.v. Mice treated with DNBSO₃ i.v. (15 mg per mouse) to induce T_S were injected i.v. with interferon or buffer 2 h later. After 6 days spleen cell suspensions were prepared and transferred by i.v. injection to naive recipients (1×10^8 cells per mouse). The recipients were sensitized with 15 μ l 0.5% DNFB and challenged as described in Fig. 1 legend. The results are expressed as mean $\pm$ s.d. from 2–3 experiments. *a*, Optimally sensitized animals. Groups *b–f*, ear swelling in recipients receiving spleen cells from DNBSO₃-treated donors: *b*, no further treatment of donors; *c*, donors injected i.v. with 0.5 ml buffer; *d–f*, donors injected i.v. with 0.5 ml interferon (1,000, 100, 10 U). *g*, Unsensitized, ear-challenged control. Statistical analysis: group *b* less than *a*; *c* less than *a*; *c* not different from *b*; *d* not different from *a*; *d* increased over *b* and *c*; *e, f* not different from *b, c*.

Figure 4 shows the results of experiments in which spleen cells from tolerized animals were treated *in vitro* with interferon before transfer to the recipient animals. The spleen cells were prepared from animals treated with DNBSO₃ 6 days earlier and 1×10^9 per ml were incubated with 10,000 U interferon per ml for 2 h at 37°C, or were incubated similarly in the control buffer or transferred immediately after preparation to the recipient animals. After incubation the cells were washed three times in Dulbecco's minimal essential medium and 1×10^8 cells per mouse injected i.v. into the recipient animals. The incubation and washing procedure of the buffer-treated spleen cells did not affect the suppressive capacity of these cells in the recipient animals (Fig. 4c) compared with the non-incubated spleen cells (Fig. 4b). Incubation of the spleen cells with interferon, however, completely abolished their capacity to transfer suppression (Fig. 4d).

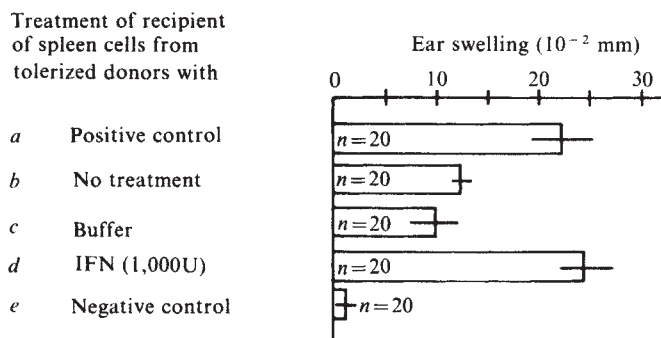


Fig. 3 Effect of interferon treatment of the recipients of T_S on the T_S response. Groups of mice were tolerized as described in Fig. 2 legend. Spleen cell suspensions were prepared after 6 days and transferred to recipients (1×10^8 cells per mouse) which were treated with buffer or interferon by i.v. injection 2 h after cell transfer. Sensitization and challenge of the recipients was performed as described in Fig. 1 legend. *a*, Optimally sensitized animals; *b–d*, ear swelling in recipients treated 2 h after T_S cell transfer with: *b*, no treatment; *c*, buffer; *d*, interferon. Statistical analysis from two to three experiments: group *b* less than *a*; *c* less than *a*; *c* not different from *b*; *d* increased over *b* and *c*; *d* not different from *a*.

The results reported here show that T_S of delayed-type hypersensitivity are inhibited by electrophoretically pure interferon; the *in vitro* treatment of T_S with interferon revealed a subpopulation of immune regulatory lymphocytes highly sensitive to interferon (1 U of interferon inhibits T_S cells present in 10^5 spleen lymphocytes). The T_{DH} response is not affected by the amount of interferon which inhibits T_S . Inhibition of T_{DH} probably requires much higher amounts as suggested by the inhibitory effects of interferon on expression of delayed-type hypersensitivity to sheep red blood cells and picryl chloride¹⁰. Interferons have been shown to affect the immune reaction *in vitro* and *in vivo* (for review see ref. 11). Interferon augments the cytotoxic activity of natural killer cells¹² and inhibits cell-mediated immune responses *in vivo*¹⁰ and *in vitro*¹³. Its effects on cell-mediated immunity, as measured in experimental contact allergy induced by picryl chloride or in delayed reaction to sheep red blood cells¹⁰, strictly depends on the interferon dose used and the timing of injection. Whereas application of interferon before sensitization results in a decreased immune

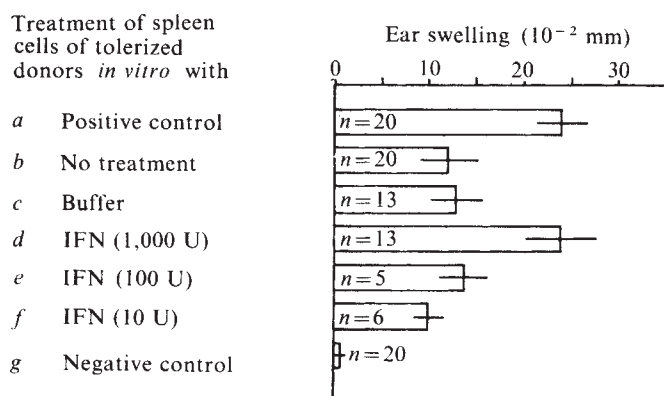


Fig. 4 Treatment of spleen cells from tolerized donors with interferon *in vitro* before transfer. Spleen cell suspensions were prepared from tolerized donors as described in Fig. 2 legend; 1×10^9 cells were incubated at 37°C for 2 h, 7% CO₂ in 1.0 ml Dulbecco's minimal essential medium containing control buffer or interferon. After incubation the cells were washed three times with the medium and injected i.v. into recipients (1×10^8 cells per mouse). Recipients were sensitized and challenged as described in Fig. 1 legend. *a*, Optimally sensitized animals; *b–f*, ear swelling in animals receiving spleen cells treated *in vitro* with: *b*, no treatment; *c*, buffer; *d–f*, interferon (1,000, 100, 10 U per 1×10^8 spleen cells). Statistical analysis of one to two experiments: groups *b, c* less than *a*; *c* not different from *b*; *d* increased over *c, b*; *d* not different from *a*; *e, f* not different from *b, c*.

response, injection after sensitization increases the immune response¹⁴. In our system which is able to differentiate between T_{DH} and T_S , small amounts of interferon were capable of increasing the immune responsiveness by selectively suppressing T_S . This probably provides an important clue to the observed enhancing effect of IFN. At the moment it is not clear which T_S (suppressing the afferent or efferent limb of the immune response) may be preferentially inhibited by interferon. The mechanism of interferon-mediated inhibition of T_S may be explained by the suppressive effect on lymphocyte proliferation which has been observed in several *in vitro* systems¹⁵. Inhibition of T_S proliferation would explain the inhibitory effect of interferon in the donor, the recipient and the *in vitro* treatment of T_S if one assumes that further clonal expansion is required for the T_S function in the recipient animals. An alternative explanation would be inhibition of effector function of T_S by interferon or an indirect action on T_S by contra suppressor cells¹⁶.

The results obtained with the purified interferon preparation confirm our previous findings that T_S inhibition by *C. parvum*² and serum obtained from *C. parvum*-treated animals³ is mediated by interferon⁴. Although it is widely accepted that interferon is immunosuppressive *in vitro* and *in vivo*, recent

observations¹⁷ showing that interferon enhanced the specific antibody response to sheep red blood cells *in vivo* and our present results suggest a stimulatory effect of interferon. The interference of interferon with T_S may augment immune responses and be an important protective mechanism against viral and bacterial infections.

We thank Maria-Theres Keuck and Reinhild Luster-Haggeney for technical assistance and Brunhilde Scheibel for typing the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (grant KN 120/4).

Received 27 November 1981; accepted 22 February 1982.

- Knop, J., Riechmann, R. & Macher, E. *J. invest. Derm.* **76**, 193–196 (1981).
- Knop, J., Riechmann, R. & Macher, E. *J. invest. Derm.* **76**, 396–399 (1981).
- Knop, J., Riechmann, R. & Macher, E. *J. invest. Derm.* **77**, 469–473 (1981).
- Knop, J., Riechmann, R. & Macher, E. *J. invest. Derm.* (submitted).
- Phanuphak, P., Moorhead, J. W. & Claman, H. N. *J. Immun.* **113**, 1230–1236 (1974).
- Sy, M. S., Miller, S. D. & Claman, H. N. *J. Immun.* **119**, 240–244 (1977).
- De Maeyer-Guignard, J., Tovey, M. G., Gresser, I. & De Maeyer, E. *Nature* **271**, 622–625 (1978).
- Tiara, H. *et al. Science* **207**, 528–529 (1980).
- Claman, H. N., Miller, S. D., Sy, M. S. & Moorhead, J. W. *Immun. Rev.* **50**, 105–132 (1980).
- De Maeyer, E., De Maeyer-Guignard, J. & Van de Putte, M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1753–1757 (1975).
- Epstein, L. B. in *Interferons and their Actions* (ed. Stewart, W. E. II) 91–132 (CRC, Cleveland, 1977).
- Zarling, J. M., Eskra, L., Borden, E. C., Horoszewicz, J. & Carter, W. *J. Immun.* **123**, 63–70 (1979).
- Heron, I., Berg, K. & Cantell, K. *J. Immun.* **117**, 1370–1373 (1976).
- De Maeyer, E. & De Maeyer-Guignard, J. *Nature* **284**, 173–175 (1980).
- Kadish, A. S., Tarsey, F. A., Yu, G. S. M., Doyle, A. T. & Bloom, B. *J. exp. Med.* **151**, 637–650 (1980).
- Gershon, R. K. *et al. J. exp. Med.* **153**, 1533–1546 (1981).
- Vignaux, F., Gresser, I. & Fridman, W. H. *Eur. J. Immun.* **10**, 767–772 (1980).

Primary structural studies of the Qa-2 alloantigen: implications for the evolution of the MHC

Mark J. Soloski, Jonathan W. Uhr & Ellen S. Vitetta

Department of Microbiology, University of Texas Health Science Center, Dallas, Texas 75235, USA

The murine major histocompatibility complex (MHC) contains a number of loci encoding cell-surface glycoproteins of molecular weight (MW) 40,000–45,000 which are non-covalently associated with β_2 -microglobulin. The products of this large multigene family, collectively referred to as class I gene products, include H-2K, -D, -L and -R, which are encoded by the H-2 complex, and Qa-1, -2 and TL, encoded by the Tla region telomeric to H-2 (refs 1–5). Functional studies have shown that the H-2K, -D and -L gene products serve as important self-recognition structures in immune responses^{6,7}; a functional role for the other class I gene products has not been established. Tryptic peptide map comparisons and primary sequence analyses have indicated that H-2K, -D and -L are 70–85% homologous with respect to their amino acid sequences^{8–10}. The data, however, have failed to reveal structural features unique to gene products of a given locus. Recent tryptic peptide map comparisons of the Tla gene products, Qa-2 and TL, have shown that Qa-2, H-2K and -D are more closely related to each other than to TL^{11,12}. Here we present the NH₂-terminal sequence of the Qa-2 alloantigen. Our results, together with those of previous studies, show that Qa-2 has significant homology to H-2 antigens but differs in that Qa-2 molecules lack extensive polymorphism and the Qa-2 heavy chain has two additional NH₂-terminal amino acids and several critical amino acid interchanges compared with H-2 antigens.

The conventional antiserum to Qa-2 alloantigens (B6.K1 α B6) is potentially polyvalent with possible reactivities to Qa-2, -3, -4 and -5 (ref. 13). However, the immunoprecipitating activity of the K1 α B6 serum is directed against the Qa-2 alloantigen as absorption of this serum with thymus (Qa-2⁺3⁺4⁺5⁺) or EL-4 lymphoma (Qa-2⁺3⁺4⁺5⁺) cells removes

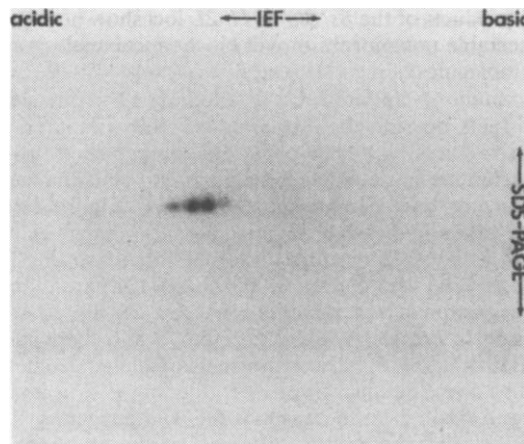


Fig. 1 Two-dimensional gel electrophoresis of the Qa-2 alloantigen. Spleen cells (5×10^7) from C57BL/6J (B6) mice were radiolabelled by lactoperoxidase-catalysed iodination³¹. Radiolabelled cells were washed with phosphate-buffered saline (PBS) and lysed in PBS containing 0.5% NP40 at 4 °C for 15 min. Nuclei and debris were removed by centrifugation and the cytoplasmic extract dialysed at 4 °C against PBS for 16 h. The dialysed extract was centrifuged at 10,000 g for 15 min to remove insoluble material before immunoprecipitation. Immunoglobulin was pre-cleared from the lysate by incubation with rabbit anti-mouse immunoglobulin (R α Mlg, 200 μ l per 5×10^7 cells) at 4 °C for 45 min, then protein A-bearing Cowan I strain of *Staphylococcus aureus* was added and the mixture incubated at 4 °C for 45 min. Immunoglobulin-depleted lysates were then treated with normal mouse serum (200 μ l per 5×10^7 cells, 4 °C, 45 min), followed by two additions of Cowan I strain of *S. aureus*. The precleared lysate was then reacted with anti-Qa-2 serum at 4 °C for 45–60 min. Immune complexes were removed by adding goat anti-mouse immunoglobulin (G α Mlg). The immunoprecipitates were washed three times with PBS containing 0.1% SDS, 0.25% NP40 and 0.2% deoxycholate followed by one wash with PBS. The washed immunoprecipitate was then analysed by two-dimensional gel electrophoresis as described by O'Farrell³² and modified by Jones³³. Gels were exposed to Fuji Rx X-ray film using a Dupont Cronex Lightning-plus intensifying screen. Exposure was for 1–2 weeks at –70 °C. IEF, immunoelectrophoresis; SDS-PAGE, SDS-polyacrylamide gel electrophoresis.

all immunoprecipitating activity¹⁴. Since it was critical for sequence studies to establish that this antiserum precipitates a single species of 40,000-MW heavy chain, Qa-2 antigens were immunoprecipitated from lysates of radioiodinated C57BL/6 spleen cells and the immunoprecipitates were analysed by two-dimensional gel electrophoresis. As shown in Fig. 1, the anti-Qa-2 antiserum precipitates several acidic 40,000-MW species. The observed microheterogeneity in Qa-2 is typical of cell-surface glycoproteins varying only in the sialic acid content of their carbohydrate moieties. Consistent with this interpretation is the observation that pretreatment with neuraminidase results in a shift to more basic species (data not shown). These results suggest that a single polypeptide chain carries the serologically defined Qa-2 determinant.

H-2 gene products have extensive polymorphism, for example, 56 alleles at the H-2K locus and 45 at the H-2D locus¹⁵. As our peptide mapping studies indicated that the Qa-2 molecules show extensive homology with H-2 gene products, it was important to determine whether Qa-2 exhibited strain-related polymorphism. So far, two Qa-2 alleles have been defined by the ability of standard antisera (B6.K1 α B6) and complement or cytotoxic effector cells either to kill (Qa-2^a) or not kill (Qa-2^b) spleen cells from various inbred mouse strains^{13,16}. Both the serological and cellular approaches have failed to detect polymorphism associated with the Qa-2^a allele. It is not clear whether the Qa-2^b allele is a true 'null' allele or whether the results are due to a reagent limitation. A failure to detect polymorphism by immunological techniques does not indicate the absence of primary structural polymorphism. Thus

the gene products of the *Ss/Slp* or *H-2L* loci show no serologically detectable polymorphism, yet biochemical analysis of the isolated molecules detects structural variations^{17,18}.

We immunoprecipitated Qa-2 alloantigens from various inbred strains bearing the *Qa-2^a* allele and compared their primary structures by tryptic peptide mapping. The sensitivity of this technique in detecting a single amino acid change in a given molecule has been well documented¹⁹. Figure 2 shows that the arginine-labelled tryptic peptide maps of Qa-2 molecules isolated from three different inbred strains (A/J, BALB/c and the wild congenic B10.KPB-128) are identical. We have examined several additional *Qa-2^a* inbred strains including B10, B6, DBA/1, SWR, DBA/2 and the wild congenic B10.KEA5; the Qa-2 molecules isolated from these strains also have identical tryptic peptide maps. Note that the background strain used to construct the wild congenic strains (B10.BR) carries the *Qa-2^b* allele^{20,21}. Thus, in those wild congenic strains typed *Qa-2^a*, the *Qa-2* locus probably derived from the wild mouse population. When the allelic products of the *H-2K* and *-D* loci were examined using similar methods, only 20–50% of their tryptic peptides were identical^{8,9}. Our observations indicate that the gene products of the *Qa-2* locus do not exhibit the extensive allelic structural variation typical of the most well characterized members (that is, *H-2K*, *-D* and *-L*) of this multigene family. Limited polymorphism may be a property of *Tla* gene products as serological studies have distinguished two alleles for the *Qa-1*, *-3*, *-4* and *-5* loci and five alleles for the *Tla* locus¹³. Thus, the present and previous peptide mapping studies present a paradox; Qa-2 is strikingly homologous with H-2 antigens, but is the first member of the class I (or class II) MHC antigens that is not highly polymorphic.

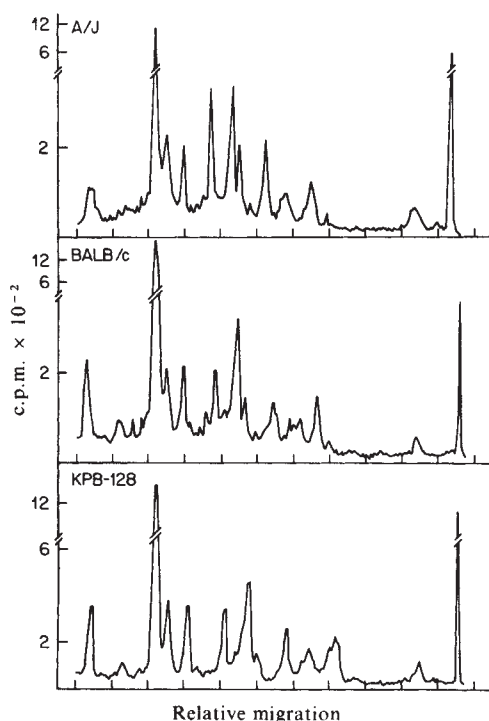


Fig. 2 Tryptic peptide maps of Qa-2 isolated from several inbred mouse strains. Qa-2 molecules metabolically labelled with ³H-arginine were isolated by immunoprecipitation and SDS-PAGE from a lysate of concanavalin A (Con A)-stimulated spleen cells¹¹. The isolated molecules were digested with trypsin and the resulting peptides analysed by cation-exchange chromatography on Technicon chromobeads type P¹¹. Before chromatography, the sample was mixed with ¹⁴C-arginine-labelled P3 myeloma κ chains, which had been digested with trypsin. These peptides served as internal markers facilitating the comparison of each peptide map. The figure shows the arginine-labelled tryptic peptide maps of Qa-2 isolated from A/J (top), BALB/c (middle) and B10.KPB-128 (bottom).

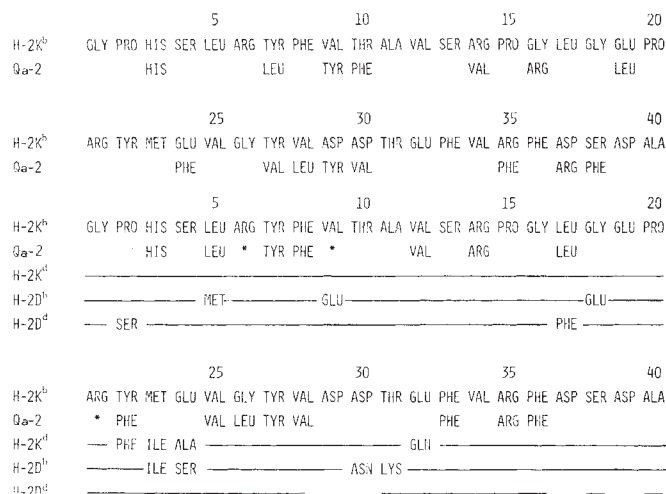


Fig. 3 Partial NH₂-terminal sequence of Qa-2 compared with other class I molecules. Qa-2 molecules metabolically labelled with a single amino acid were immunoprecipitated from extracts prepared from Con A-stimulated B6 spleen cells. The immunoprecipitates were denatured, reduced and the 40,000-MW Qa-2 molecule isolated by SDS-PAGE¹¹. The isolated Qa-2 molecules were lyophilized, re-dissolved in H₂O with 1 mg human γ -globulin as carrier, and dialysed against H₂O to remove free SDS. The dialysed Qa-2 molecule was loaded on to a Beckman 890C sequencer for analysis^{22,23}. In each sequencing run, a ³⁵S-methionine-labelled light chain was used as an internal standard. The methionine residues at positions 4, 11 and 13 allowed the determination of a repetitive yield. Thiazolinone derivatives were evaporated under nitrogen and counted in a liquid scintillation counter. Assignments were based on: (1) the appearance of a peak of radioactivity in a given sequencing step; (2) the ability of multiple peaks to fit a standard repetitive yield curve as determined for each run using the ³⁵S-methionine immunoglobulin light-chain standard. The H-2 sequences were obtained from Coligan *et al.*¹⁰. The upper panel shows the non-aligned H-2K^b and Qa-2 assignments. The lower panel shows the Qa-2 assignments aligned with several H-2 sequences after the introduction of a 2-amino acid gap in the Qa-2 sequence after position 3. Asterisks in the lower panel denote the positions of valine and arginine residues in H-2 molecules, and their counterparts in other species, which are absent from the Qa-2 molecule. The solid lines indicate residues which are identical to those in H-2K^b.

To elucidate the primary structure of the Qa-2 alloantigen, we analysed the NH₂-terminal sequence. Qa-2 alloantigen labelled with a single ³H-amino acid was isolated from lysates of mitogen-activated B6 splenocytes which have been shown to be particularly enriched for the expression and synthesis of Qa-2 (ref. 11). The 40,000-MW heavy chain of Qa-2 was prepared by electrophoresis of the immunoprecipitate and used for NH₂-terminal sequencing^{22,23}. Figure 3 (top) shows the partial NH₂-terminal sequence of Qa-2 in which we have determined the Arg, Leu, His, Phe, Val and Tyr residues within the first 40 amino acids. The sequence of the H-2K^b molecule is also shown. Surprisingly, only 1 of 15 assignments show identity. However, if we assume that during evolution two additional amino acids have been inserted between positions 3 and 7, then shifting the Qa-2 assignments two positions towards the NH₂-terminus gives 13/15 assignments which are identical with their H-2K^b counterparts (Fig. 3, bottom). This level of homology confirms the primary structural homology predicted from tryptic peptide map comparisons of Qa-2 and H-2K or -D molecules¹¹. Further interesting features are: (1) the location of two additional amino acids of Qa-2 on the C-terminal side of position 3 and on the NH₂-terminal side of position 7; (2) primary sequence analysis of murine H-2K, -D and -L and their counterparts in other species has shown that the arginines at positions 5, 14, 21 and 35 are highly conserved residues²⁴. Qa-2, on the other hand, has been assigned arginines only at positions 14 and 35 (hereafter we shall refer to the Qa-2 assignments obtained after the 2-amino acid shift). Furthermore, H-2

homologues also show a highly conserved glycine at position 26 (ref. 24); Qa-2 has been assigned a leucine at that position. In addition, Qa-2 lacks a valine at position 9 that is found in H-2K^b, -K^d and -D^b molecules. Comparison of the 18 residues (12 assignments, the absence of a valine residue at position 9, and the absence of two arginine residues at 5 and 21) indicates 72–82% homology between Qa-2 and H-2K^b, -K^d, -D^b or -D^d. H-2K and -D molecules show 83–94% homology for these 18 positions. These observations indicate that Qa-2 lacks several highly conserved primary structural features of MHC-controlled major transplantation antigens. Moreover, the Qa-2 alloantigen is rather less related to H-2 alloantigens than they are to each other.

The data therefore suggest that Qa-2 is part of a subfamily of class I gene products, the members of which are distinct at the primary structural and, perhaps, functional levels from members of a second subfamily which includes H-2K, -D and -L. This hypothesis is consistent with the difference between Qa-2 and H-2 antigens with regard to tissue distribution, ontogeny, role as self-recognition units and particularly, extent of polymorphism¹³. Several investigators have postulated that the evolution of polymorphism at H-2 loci was critical for the survival of species/populations against infectious agents^{15,25}. The failure of the Qa-2 locus to display extensive polymorphism may indicate a different function for the Qa-2 molecule. Interestingly, tryptic peptide map comparisons of TL with H-2K, -D and Qa-2 indicate that TL is only distantly related to these molecules, which suggests that there may be a third subfamily containing TL (refs 12, 26 and M.J.S., K. McIntyre, J.W.U. and E.S.V., in preparation).

The multigene family encoding class I gene products possibly evolved from a primordial gene that has undergone duplication and subsequent divergence²⁷. The unique primary structural features of Qa-2 suggest that it evolved before H-2K, -D and -L. Furthermore, TL may have diverged before Qa-2 because of the observed modest primary structural homology of TL with Qa-2, H-2K or H-2D (refs 12, 26 and M.J.S. *et al.*, in preparation). Indeed, the unique structural features of the Qa-2 molecule and the finding of Qa and TL homologies in other species perhaps indicate that the Qa-2 and TL duplications occurred before speciation^{28–30}. Alternatively, the multiple class I loci may have duplicated at approximately the same time and the observed structural differences evolved as a result of different selective pressures on the individual loci.

We thank Drs Mark Siegelman and J. Donald Capra for advice, discussion and technical support during these studies; Ms Patricia Wheeler and Dr Brad Ozanne for assistance in the two-dimensional gel analysis; Ms Sandy Graham, Shirley Shanahan, Kay Snavely and Martha Miles for technical assistance; and Ms Daisy Marcoulides for typing the manuscript. This work was supported by NIH grants AI-13448 and AI-11851.

Received 14 January; accepted 3 March 1982.

- Vitetta, E. S. & Capra, J. D. *Adv. Immun.* **26**, 147–193 (1978).
- Hansen, T. *et al. J. Immun.* **126**, 1713–1716 (1981).
- Stanton, T. H. & Hood, L. *Immunogenetics* **11**, 309–314 (1980).
- Cook, R., Rich, R. & Flaherty, L. *J. Immun.* **127**, 1894–1898 (1981).
- Michaelson, J., Flaherty, L., Vitetta, E. & Poulik, M. *J. exp. Med.* **145**, 1066–1070 (1977).
- Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 52–177 (1979).
- Zinkernagel, R. M. *Immun. Rev.* **42**, 224–270 (1978).
- Brown, J. L., Kato, K., Silver, J. & Nathenson, S. G. *Biochemistry* **13**, 3174–3178 (1974).
- Sears, D. W., Young, S., Wilson, P. H. & Christiaansen, J. *J. Immun.* **124**, 2641–2649 (1980).
- Coligan, J. E., Kindt, T. J., Uehara, H., Martinko, J. & Nathenson, S. G. *Nature* **291**, 35–39 (1981).
- Soloski, M. J., Uhr, J. W., Flaherty, L. & Vitetta, E. S. *J. exp. Med.* **153**, 1080–1093 (1981).
- McIntyre, K., Hämmerling, U., Uhr, J. W. & Vitetta, E. S. *J. Immun.* (in the press).
- Flaherty, L. in *The Role of the Major Histocompatibility Complex in Immunobiology* (ed. Dorf, M.) 33–57 (Garland, New York, 1981).
- Michaelson, J., Flaherty, L., Bushkin, Y. & Yudkowitz, H. *Immunogenetics* **14**, 129–140 (1981).
- Klein, J. *Science* **203**, 516–521 (1979).
- Forman, J. & Flaherty, L. *Immunogenetics* **6**, 227–233 (1978).
- Parker, L. L., Atkinson, J. P., Roos, M. H. & Shreffler, D. C. *Immunogenetics* **11**, 55–64 (1980).
- Rose, S., Hansen, T. & Cullen, S. *J. Immun.* **125**, 2044–2049 (1980).
- Nairn, R., Yamaga, Y. & Nathenson, S. G. *A. Rev. Genet.* **14**, 241–277 (1980).
- Zaleska-Rutczynska, S. & Klein, J. *J. Immun.* **119**, 1903–1911 (1977).
- Klein, J. *Transplantation* **13**, 291–299 (1972).
- Cook, R. G. *et al. J. Immun.* **121**, 1015–1019 (1978).

- Cook, R. G., Siegelman, M. S., Capra, J. D., Uhr, J. W. & Vitetta, E. S. *J. Immun.* **122**, 2232–2237 (1979).
- Ploegh, H., Orr, H. & Strominger, J. *Cell* **24**, 287–299 (1981).
- Zinkernagel, R. M. *A. Rev. Microbiol.* **33**, 201–213 (1979).
- Yokoyama, K., Stockert, E., Old, L. J. & Nathenson, S. G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7078–7082 (1981).
- Maizels, R. M., Fellinger, J. A. & Hood, L. *Immunogenetics* **7**, 425–444 (1978).
- Schwartz, B. D., Ciceri, R., Berggard, I. & Schevach, E. *J. Immun.* **121**, 835–839 (1978).
- Gazit, E., Terhorst, C., Mahoney, R. & Yunis, E. *Hum. Immun.* **1**, 97–109 (1980).
- Terhorst, C. *et al. Cell* **23**, 771–780 (1981).
- Baur, S., Vitetta, E. S., Sherr, C. J., Schenkein, I. & Uhr, J. W. *J. Immun.* **106**, 1133–1135 (1971).
- O'Farrell, P. *J. biol. Chem.* **250**, 4007–4021 (1975).
- Jones, P. *J. exp. Med.* **146**, 1261–1279 (1977).

Mouse immunoglobulin allotypes: post-duplication divergence of $\gamma 2a$ and $\gamma 2b$ chain genes

Roger Ollo & Francois Rougeon

Unité de Génétique et Biochimie du Développement, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cedex 15, France

Sequence analysis of the mouse immunoglobulin heavy chain constant-region gene family has provided interesting clues to its evolutionary history¹. There are eight genes, each consisting of four exons separated by noncoding sequences, which encode the three globular domains and the small hinge region of the constant-region polypeptide. The eight genes are believed to have arisen by duplications from a single ancestral gene and in particular, the $\gamma 2a$ and $\gamma 2b$ genes are thought to have been the result of a recent duplication^{2–5}. They show, however, a rather surprising pattern of sequence divergence^{2–7}. Schreier *et al.*⁷ have reported considerable sequence divergence between the two $\gamma 2a$ alleles carried by the BALB/c and C57BL/6 inbred mouse strains; at many of the sites at which the two alleles differ, the BALB/c allele is identical to the BALB/c $\gamma 2b$ gene. Thus it has been suggested that the BALB/c $\gamma 2a$ gene has been converted by the $\gamma 2b$ gene. We have now sequenced the C57BL/6 $\gamma 2b$ gene and report that the divergence between the two $\gamma 2b$ alleles is very much less than that between the $\gamma 2a$ alleles and is consistent with point mutations. However, it is also clear that the third domain (CH3) of the C57BL/6 $\gamma 2a$ gene has diverged so much from the homologous domain in the BALB/c $\gamma 2a$ gene and the two $\gamma 2b$ alleles that it may have derived from elsewhere in the heavy chain family and become inserted in the $\gamma 2a$ gene by a double unequal crossing-over event.

The complete nucleotide sequence of the $\gamma 2b^b$ chain C gene, determined as described in Fig. 1 legend, is shown in Fig. 2. A total of 1,922 nucleotides were determined. The nucleotide sequence predicts the complete amino acid (336 residues) sequence encoded by the $\gamma 2b^b$ allele. We have compared the nucleotide sequence of the $\gamma 2b^b$ allele reported elsewhere⁸ with that of the $\gamma 2b^b$ allele for over 1,834 positions. As expected for alleles at a single genetic locus, the two sequences exhibit a high level of homology; the pattern of homology is given in Table 1. In the protein-coding regions, the two sequences differ by seven base transitions. There are four G→A transitions which lead to four amino acid substitutions: two located in the CH2 domain [A→G at position 1,083 (Gln→Arg) and A→G at position 1,175 (Thr→Ala)] and two located in the CH3 domain [A→G at position 1,604 (Asn→Asp) and G→A at position 1,609 (Met→Ile)]. There are three silent substitutions resulting from T→C transitions, one in each domain. Curiously, one of these is located adjacent to the A→G transition, leading to the amino acid change Thr→Ala in the CH2 domain. It is highly probable that the double transition CA→TG leading to the Thr→Ala change in the CH2 domain is the result of a double mutation in the $\gamma 2b^a$ gene as the same TG is found in the two $\gamma 2a$ alleles.

It is generally assumed that intervening sequences (IVSs) and flanking sequences of a gene evolve more rapidly than the

Table 1 Comparison of nucleotide sequences of BALB/c and C57BL/6 $\gamma 2b$ and $\gamma 2a$ constant-region alleles

$\gamma 2b^b/\gamma 2b^a$				$\gamma 2a^b/\gamma 2a^a$			
Gene segment	No. of transitions	No. of transversions	Deletions/insertions	Gene segment	No. of transitions	No. of transversions	Deletions/insertions
CH1	1	-	-/-	CH1	9	3	-/-
H	-	-	-/-	H	3	8	-/15
CH2	3	-	-/-	CH2	6	5	-/-
CH3	3	-	-/-	CH3	18	28	3/3
3'UT	-	-	-/-	3'UT	8	7	-/-

protein-coding sequence^{9,10}. It is interesting to note that the noncoding regions of both alleles are highly conserved with respect to size and sequence. There are only three base changes located in IVS1 and IVS2 and one in the 3' genomic flanking region, while the 3'-untranslated region is unchanged. Furthermore, the base deletion located in IVS1 is followed 16 nucleotides downstream by a base insertion, leaving the length of the IVS1 identical in both alleles.

The genetic polymorphism of mouse immunoglobulin in heavy chains has been extensively studied using anti-allotype allosera^{11,12}. However, in only a few cases have the structural basis and the precise location of these polymorphic determinants been determined¹³. By extensive specificity mapping, Herzenberg *et al.*¹¹ characterized six alleles at the $\gamma 2b$ locus, the BALB/c $\gamma 2b^a$ allele and the C57BL/6 $\gamma 2b^b$ allele differing by three antigenic specificities. Because of limited amino acid variations between the products of both alleles, it is interesting to correlate the four amino acid differences to allotypic determinants. The first two amino acid changes (Glu→Arg, Thr→Ala), located in the CH2 domain, are separated by 31 amino acids, and thus are likely to determine separate specificities. A Thr→Ala change also occurs in the same region of rabbit γ -chain, allowing the correlation of the A14 and A15 specificities with this amino acid change¹⁴. As the last two amino

acid changes are contiguous, it is reasonable to correlate the third allotypic specificity to the double change Asn-Met→Asp-Ile (positions 1,604 and 1,609).

From our data and those of Schreier *et al.*⁷, it appears that the two closely linked $\gamma 2b$ and $\gamma 2a$ chain genes have different rates of evolution (Table 1). Whereas the two $\gamma 2b$ chain alleles differ by only 12 single base changes in 1,850 positions, the two $\gamma 2a$ alleles differ in 111 of 1,093 positions compared. The observation that some of the nucleotides that differ between the two $\gamma 2a$ alleles are conserved between $\gamma 2b$ and $\gamma 2a$ chain genes in BALB/c mice led Schreier *et al.*⁷ to postulate an intergenic conversion between the two genes. Our data show that these nucleotides are conserved between $\gamma 2a^a$, $\gamma 2b^a$ and $\gamma 2b^b$ genes. Furthermore, these positions are scattered along the sequence and not clustered as in the human fetal γ -globin genes¹⁵. Therefore, if there has been any gene conversion, it could have occurred only in the BALB/c ancestor.

The mutation rate of the $\gamma 2b$ chain gene is as expected for single base mutations in two alleles. However, the mutation rate in the $\gamma 2a$ gene, which is 20 times higher than that in the $\gamma 2b$ gene, is too great to be explained in this way. Note also that most of the differences between $\gamma 2a$ alleles are localized in the CH3 domain. Previous comparative studies of $\gamma 2a$ and $\gamma 2b$ genes led to the hypothesis that segments of the γ genes

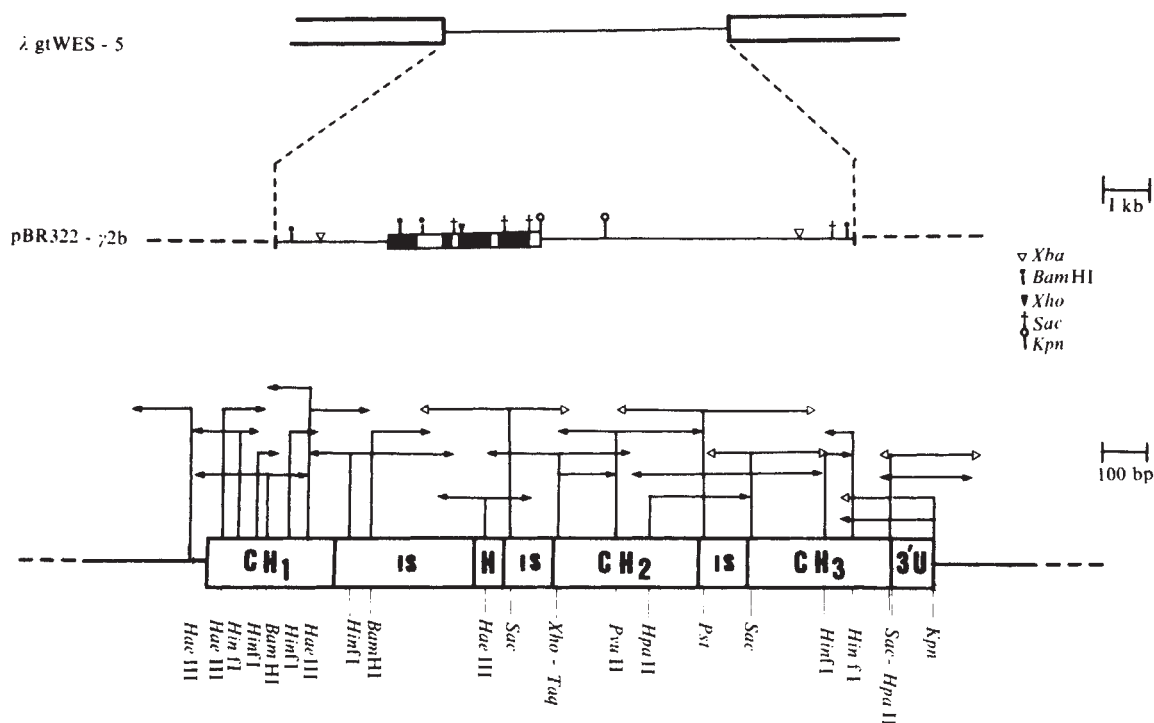


Fig. 1 Sequencing strategy for the mouse $\gamma 2b$ chain constant-region gene. As the $\gamma 2b$ gene lies on a 6.6 kilobase (kb) *Eco*RI fragment in C57BL/6 mouse embryo DNA¹⁶, *Eco*RI fragments (fractionated by sucrose centrifugation) ranging in size from 6 to 8 kb were ligated with λ gtWES arms purified by gel electrophoresis¹⁷ and packaged *in vitro*. A hybridization probe consisting of a *Xba*-*Kpn* fragment of the BALB/c $\gamma 2b$ gene cloned in phage Charon 4A (unpublished results) was used to isolate a recombinant phage λ gtWES- $\gamma 2b$ C57. For sequence analysis, the *Eco*RI insert was transformed to pBR322. A restriction enzyme map of this insert in pBR322 was determined by single and double digestion of DNA with various enzymes. The strategy used for sequence determination is shown at the bottom. The extent and direction of sequencing of each fragment are indicated by the horizontal arrows: open and solid arrows represent 3' and 5' end-labelling, respectively. Only the restriction sites used as starting points are shown.

[illegible]

© 1982 Macmillan Journals Ltd

Electrophoresis was conducted on Cellogel², and aconitase made visible by the method of Harris and Hopkinson³. Using this system, two zones of aconitase activity are found in bird liver. By differential centrifugation of mitochondrial and cytoplasmic fractions it was shown that the more anodally migrating zone of activity was the cytoplasmic fraction (Fig. 1). The cytoplasmic isozyme of aconitase (Acon_c) from liver of the guinea fowl (*Numida meleagris*—order Galliformes) migrates more anodally than that of domestic fowl (*Gallus domesticus*—order Galliformes) (Fig. 1). We denote the domestic fowl form of Acon_c as A and that of the guinea fowl as B. Hybrids between guinea fowl and domestic fowl were produced by artificial insemination⁴.

If the structural gene for Acon_c were autosomal, all offspring from such a cross would be heterozygous AB. If, however, the locus for Acon_c were on the Z chromosome, male offspring would be heterozygous AB whereas female offspring would be hemizygous A or B depending on the direction of the cross.

Crosses between female guinea fowl and male domestic fowl yielded hybrid embryos most of which died within the first week of incubation. A small number survived for longer periods, including one that survived to hatching. A total of 17 embryos were successfully typed for aconitase and of these, five were of known sex either from gonad histology or karyotyping⁵. Seven of the embryos were AB but ten were A (Table 1). Moreover,

Table 1 Cytoplasmic aconitase types of hybrid embryos between female guinea fowl (type B) and male domestic fowl (type A)

	Acon _c type		
	A	AB	B
Male embryos	0	3	0
Female embryos	2	0	0
Unsexed embryos	8	4	0
Total	10	7	0

of the five embryos of known sex, the three males were AB and the two females were A (Table 1 and Fig. 1). The results therefore indicate that the locus for Acon_c is not autosomal in the guinea fowl, but is located on the Z chromosome.

The data do not exclude the possibility that Acon_c is autosomally inherited in domestic fowl and sex-linked in guinea fowl. Crosses between female domestic fowls and male guinea fowls yielded embryos that died within 24 h. Aconitase activity was undetectable in most of these embryos. However, three embryos yielded a faint band of aconitase activity in the position of guinea fowl Acon_c and no detectable activity in the position of domestic fowl Acon_c. These results are expected for female hybrid embryos if Acon_c is encoded by a locus on the Z chromosome of the domestic fowl, and suggest therefore that Acon_c is Z-linked in the domestic fowl also.

We have found that Acon_c is polymorphic in three species of birds—the little corella, *Cacatua sanguinea* (order Psittaciformes), the yellow-tailed black cockatoo, *Calyptrorhynchus funereus* (order Psittaciformes), and the house sparrow *Passer domesticus* (order Passeriformes).

Table 2 shows the observed phenotype frequencies for Acon_c in little corellas from Three Springs, Western Australia. Of five males, two were type F (fast migrating band) and three were type FS (double banded fast and slow), whereas of 12 females, nine were type F and three were type S (Table 2). No heterozygous females were found. The appropriate statistical test is to construct a 2×2 contingency table for heterozygote and homozygote males versus females. Applying Fisher's exact probability, *P* is 1.5%. The absence of heterozygous females is therefore statistically significant. If we take the gene frequencies for males and calculate expected phenotype frequencies for females on the hypothesis that Acon_c is sex-linked, the fit to the observed results is then excellent (Table 2), suggesting that the locus for Acon_c is on the Z chromosome of *C. sanguinea*.

In five yellow-tailed black cockatoos (two males and three females) from Bool Lagoon, South Australia, three alleles for Acon_c were found, designated F, I and S (fast, intermediate and slow). One male was heterozygous FI, and the other was heterozygous IS, but of the three females, one was single-banded F, one was single-banded I and the third was single-banded S. Again these data are more consistent with the locus for Acon_c being on the Z chromosomes of *C. funereus* than on an autosome.

Samples of house sparrows (*P. domesticus*) from Adelaide, Canada and the UK were also polymorphic for Acon_c (data for Canada and UK from S. R. Cole and D. Parkin, personal communication). Again there was an absence of heterozygous females which is highly significant statistically in the UK sample, and significant at the 5% level in the Canadian sample (Table 2). The observed phenotype distributions for females are, on the other hand, consistent with the locus for Acon_c being on the Z chromosome of *P. domesticus*.

Ohno¹ has suggested that the Z chromosome has been conserved during evolution of the class Aves; thus any gene that is on the Z chromosome of one species of bird should be found on the Z chromosome of all bird species. The present study has shown that Acon_c is on the Z chromosome of a galliform and probably on the Z chromosome of another galliform, two psittaciforms and a passeriform. These data provide support for Ohno's hypothesis and suggest that Acon_c is on the Z chromosome of all birds. The data do not prove, however, that the entire Z chromosome has been conserved, rather that part of the Z chromosome carrying the structural locus for Acon_c has been conserved.

Ohno¹ also suggested that the Z chromosome of birds may be homologous with the X chromosome of mammals. Clearly, this is not the case. The enzymes glucose-6-phosphate dehydrogenase and phosphoglycerate kinase are sex-linked in

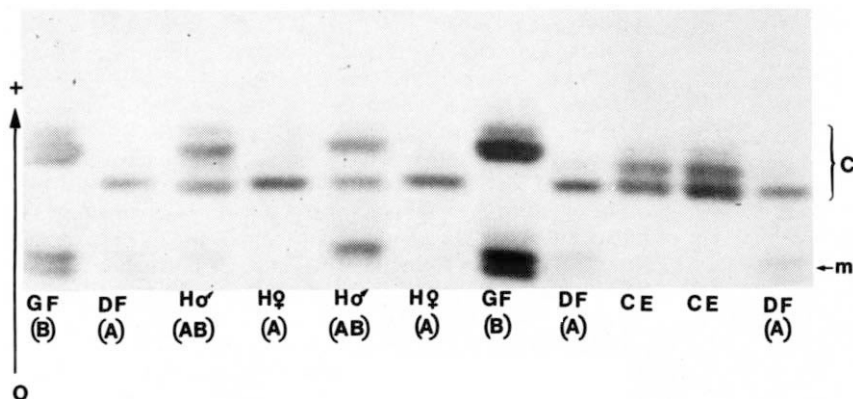


Fig. 1 Electrophoretic gel stained for aconitase. O, origin; m, mitochondrial fraction (Acon_m); C, cytoplasmic fraction (Acon_c); GF, guinea fowl (type B for Acon_c); DF, domestic fowl (type A for Acon_c); H, hybrid between female guinea fowl and male domestic fowl; CE, cytoplasmic extract of chicken liver after differential centrifugation to remove the mitochondrial fraction. Note that the proposed heterozygotes are double-banded, indicating that avian cytoplasmic aconitase is monomeric as in mammals³. The weaker secondary bands result from post-translational modification.

Table 2 Phenotype frequencies for Acon_c by sex in little corellas from Three Springs, Western Australia, and in house sparrows from UK, Canada and Adelaide

a, Little corella							c, House sparrows—Canada						
			F	FS	S	Total				F	FS	S	Total
<i>p</i> = 0.7 <i>q</i> = 0.3	<i>Observed:</i>	Male	2	3	0	5	<i>p</i> = 0.167 <i>q</i> = 0.833	<i>Observed:</i>	Male	2	9	28	39
		Female	9	0	3	12			Female	5	0	16	21
	<i>Expected female:</i>	Autosomal	5.9	5.0	1.1	<i>Expected female:</i>		Autosomal	0.6	5.8	14.6		
		Sex-linked	8.5	0	3.6			Sex-linked	3.5	0	17.5		
				Hetero-zygote	Homo-zygote	Total					Hetero-zygote	Homo-zygote	Total
			Male	3	2	5				Male	9	30	39
		Female	0	12	12			Female	0	21	21		
			3	14	17				9	51	60		
		<i>P</i> = 1.5% (exact test)							<i>P</i> = 1.5% (exact test)				

b, House sparrows—UK							d, House sparrows—Adelaide						
			F	FS	S	Total				F	FS	S	Total
<i>p</i> = 0.086 <i>q</i> = 0.914	<i>Observed:</i>	Male	0	39	188	227	<i>p</i> = 0.107 <i>q</i> = 0.893	<i>Observed:</i>	Male	0	3	11	14
		Female	18	0	145	163			Female	1	0	18	19
	<i>Expected female:</i>	Autosomal	1.21	25.62	136.2	<i>Expected female:</i>		Autosomal	0.2	3.6	15.2		
		Sex-linked	14.00	0	149.0			Sex-linked	2.0	0	17.0		
				Hetero-zygote	Homo-zygote	Total					Hetero-zygote	Homo-zygote	Total
			Male	39	188	227				Male	3	11	14
		Female	0	163	163			Female	0	19	19		
			39	351	390				3	30	33		
		<i>P</i> < 0.01% (exact test)							<i>P</i> = 6.7% (exact test)				

mammals but not birds⁶. Acon_c, on the other hand, is sex-linked in birds, but not mammals³.

Because Acon_c is Z-linked in birds, males have two doses of the gene and females one dose. Therefore in the absence of dosage compensation, male birds will possess twice the activity of cytoplasmic aconitase of females. We have measured the level of cytoplasmic aconitase activity in the livers of 1-day-old chickens, mature house sparrows and mature spotted turtle-doves (*Streptopelia chinensis*—order Columbiformes). In all three species, females had statistically significantly lower levels of cytoplasmic aconitase activity than males (Table 3). On the

other hand, where the null hypothesis was tested that males possess twice the activity of females, none of the *P* values were statistically significant (Table 3). That such a result was not due to generally lower enzyme activity in females was shown by comparing the levels of isocitrate dehydrogenase (IDH) activity in the same cytoplasmic fractions. In all three species, males and females did not differ significantly with respect to IDH activity. It is clear, therefore, that the locus for Acon_c does not show total dosage compensation in birds.

Dosage compensation of all but one or two X-linked genes in mammals is effected by inactivation in each cell of one of

Table 3 Aconitase activity in the supernatant of liver homogenates of males and females of three species of birds representing three orders

Species	Sex	<i>N</i>	Aconitase activity	Null hypothesis (1) $\mu_1 = \mu_2$			Null hypothesis (2) $\mu_1 = 2\mu_2$		
				<i>t</i> *	d.f.	<i>P</i>	<i>t</i>	d.f.	<i>P</i>
<i>P. domesticus</i> (Passeriformes)	F	5	0.095 ± 0.023	3.11	7.2	<1%	2.58	4.4	>5%
	M	5	0.135 ± 0.016						
<i>G. domesticus</i> (Galliformes)	F	4	0.032 ± 0.005	3.01	5.7	<2.5%	1.84	6.9	>20%
	M	5	0.051 ± 0.012						
<i>S. chinensis</i> (Columbiformes)	F	4	0.034 ± 0.004	3.67	7.9	<0.5%	1.24	9.6	>20%
	M	8	0.081 ± 0.025						

Livers were removed from freshly killed individuals and placed in twice their weight of cold homogenizing solution (0.25 M sucrose, 0.001 M EDTA in 0.1 M Tris-HCl pH 7.8). The tissues were homogenized by hand and centrifuged at 10,000g at 4°C for 25 min. The supernatant was removed and stored frozen at -80°C. Electrophoresis of the supernatant revealed only a trace of the mitochondrial isozyme of aconitase (Fig. 1). Aconitase activity in the supernatant was assayed by a modification of the method of Ruse and O'Connell¹⁰, isocitrate dehydrogenase activity by the method of Cleland¹¹, and total protein by the method of Lowry¹². Values are mean ± s.d. μ mol NADP converted per min per mg protein. *N*, number of individuals.

* *t*-Tests were conducted by comparing activities of cytoplasmic aconitase on the null hypotheses (1) that males and females had the same activity and (2) that males had twice the activity of females. For all tests, the degrees of freedom were calculated for variances not assumed to be equal.

the X chromosomes⁷. In *Drosophila* most genes on the X chromosome are dosage-compensated, but here each gene on the X chromosome produces half the amount of product in females as in males⁸. However, not all genes on the X chromosome of *Drosophila* show complete dosage compensation. Therefore the observation that *Acon*_c shows no dosage compensation in birds should not be taken as proof that all Z-linked genes in birds show no dosage compensation. Nevertheless, when taken in conjunction with the morphological data¹, the present data suggest that lack of dosage compensation is the rule in birds. As pointed out by Johnson and Turner⁹, it may be no coincidence that the groups known to show dosage compensation (mammals and *Drosophila*) have XX/XY sex determination while the groups lacking dosage compensation (lepidopterans and birds) have ZZ/ZW sex determination.

For their help with various aspects of this work we thank Drs D. Briscoe, R. Holmes, C. H. S. Watts, B. J. Richardson, D. Hayman, D. Saunders, G. Smith, D. Parkin, R. Hope, Ms C. Lynch, Messrs T. Benson, J. Bourne and S. R. Cole.

Received 7 October 1981; accepted 5 March 1982.

1. Ohno, S. *Sex Chromosomes and Sex-linked Genes* (Springer, New York, 1967).
2. Baverstock, P. R., Watts, C. H. S., Adams, M. & Gelder, M. *Aust. J. Zool.* **28**, 553-574 (1980).
3. Harris, H. & Hopkinson, D. A. *Handbook of Enzyme Electrophoresis in Human Genetics* (North-Holland, Amsterdam, 1976).
4. Lake, P. E. & Stewart, S. M. *Ministry of Agriculture, Fisheries & Food Bulletin* **213** (1978).
5. Fehchheimer, N. S. & Jaffe, W. P. *Nature* **211**, 773-774 (1966).
6. Cam, A. E. & Cooper, D. W. *Biochem. Genet.* **16**, 261-270 (1978).
7. Lyon, M. F. *Biol. Rev.* **47**, 1-35 (1972).
8. Lucchesi, J. C. A. *Rev. Genet.* **7**, 225-237 (1973).
9. Johnson, M. S. & Turner, J. R. G. *Heredity* **43**, 71-77 (1979).
10. Ruse, I. A. & O'Connell, E. L. *J. biol. Chem.* **242**, 1870-1879 (1967).
11. Cleland, W. W., Thompson, V. W. & Barden, R. E. *Meth. Enzym.* **13**, 30-32 (1969).
12. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).

Evidence for X-linkage and non-inactivation of steroid sulphotase locus in wood lemming

H-H. Ropers

Institut für Humangenetik und Anthropologie, Albertstrasse 11,
D-7800 Freiburg i. Br., FRG

U. Wiberg

Institute of Genetics, University of Lund, Sölvegatan 29,
S-223 62 Lund, Sweden

One of the two X chromosomes is inactivated in somatic cells of adult female mammals to compensate for unequal amounts of X-chromosomal genes in the two sexes¹. Although the general validity of this concept is not in doubt, there is evidence that a segment of the distal short arm of the human X chromosome carrying the gene, or genes for the Xg blood group system, the steroid sulphotase (STS) locus^{2,3}, and a gene controlling a serologically defined male-specific antigen^{4,5}, is never inactivated. It is not known whether this non-inactivated segment is a special feature of the human X chromosome or whether it is a general feature of the ancestral, highly conserved, mammalian X chromosome⁶. In man the number of functional STS gene copies can be deduced from intracellular STS activity⁷. With this in mind we have investigated the number of active gene copies on the X chromosome of the wood lemming (*Myopus schisticolor*, Lilljeborg), by assaying STS in cultured cells of this species. We report here that STS activity is directly correlated to the number of X chromosomes and unrelated to the phenotypic sex. This suggests that in the wood lemming the STS gene is also X-linked and not subject to inactivation.

The remarkable stability of the mammalian X chromosome can be explained, in part, by the fact that males carrying X-autosomal translocations are mostly sterile. This is thought to result from a failure of inactivation of translocated X-chromosome segments during gametogenesis⁸. As a result, X-autosomal translocations cannot be transmitted through males, and therefore do not spread in the population. This notion suggests that strict evolutionary conservation may be confined to those parts of the X chromosome that are subject to inactivation, as males with translocations involving permanently active segments would not necessarily be sterile. In man the existence of a non-inactivated segment on the distal short arm of the X chromosome is now well established⁹. For one of the three genes assigned to this region (that is, the STS gene), in both sexes, significant differences in enzyme activity have been demonstrated for various cells and tissues such as fibroblasts, hair roots and placenta^{7,10,11}. Similar gene-dosage relationships have been observed in numerical, and most structural, X-chromosomal aberrations¹².

In the wood lemming, numerical sex chromosome aberrations are relatively frequent due to its unique system of sex determination. There are two different types of X chromosome, X and X*, which can be distinguished cytogenetically¹³. Specimens with XX, X*X, X*Y, XO and X*O sex chromosome constitutions are phenotypic females and are fertile. Males have the XY sex chromosomes^{13,14}. In germ-line cells of X*Y females only the X* chromosome is found. By a mechanism of mitotic double non-disjunction, the Y chromosome is replaced by a copy of the X* (refs 15, 16). The relatively high incidence of numerical sex chromosome aberrations probably depends on occasional failure of this non-disjunction mechanism.

The molecular basis of sex reversal in X*Y females is not yet understood. The fact that the X* differs from the X both morphologically and functionally suggests that this may be due to duplication or deletion of genes involved in sex differentiation. In man, a gene or genes controlling sex differentiation may be closely linked to the STS locus on the non-inactivated distal part of Xp^{4,5,17,18}. To determine whether in a rodent species the STS gene is also X-linked and not inactivated, and whether the dosage of this gene is altered on the rearranged X*, we have determined STS activity in cultured fibroblasts from wood lemmings of seven different sex chromosome constitutions.

Primary fibroblast cultures were established from heart tissues in Earle's modified Eagle's minimal essential medium supplemented with 15% fetal calf serum. After karyotyping and two or three further passages in culture, the cells were analysed biochemically. In three subsequent series of experiments, a total of 54 cultures, representing 27 wood lemming specimens, were tested for STS activity; 30 cultures were also analysed for an X-linked reference enzyme, that is, glucose-6-phosphate dehydrogenase (G6PD).

STS activities were found to be directly correlated with the number of X chromosomes, irrespective of the phenotypic sex (Table 1) while no such correlation was found for G6PD (Table 2). The simplest interpretation of these results is that in the wood lemming, as well as in man, STS is X-linked and not subject to X-inactivation. Therefore, we suggest that the STS gene may be part of the conserved X-chromosomal linkage group of mammals, although this has yet to be confirmed in other mammalian species. Conservation of a permanently active segment of the mammalian X chromosome could be related to the hypothesized homology between this segment and part of the Y chromosome^{19,20} for which there is now increasing, although circumstantial, evidence^{17,18,21}. It is noteworthy that in man, genes on this segment have been implicated in sex determination^{4,5,17,18}, suggesting that sex reversal in the wood lemming may be due to altered dosage of related genes on the rearranged X*. While this possibility cannot be excluded, the observation that STS levels are almost identical in wood lemming specimens, whether they carry the normal X or the altered

Table 1 Steroid sulphatase activity in cultured fibroblasts of wood lemmings with different sex-chromosome constitutions

Animal no.	Phenotypic sex	Sex chromosomes	Expt 1 No. of X chromosomes		Expt 2 No. of X chromosomes		Expt 3 No. of X chromosomes	
			1	2	1	2	1	2
593	♂	XY						
598	♂	XY	240		206*			
4,849	♂	XY			355		279	
4,988	♂	XY					446	
4,360	♂	XY					442	
4,992	♂	XY					390	
4,656	♂	XY					434	
597	♀	XX		563		827		
600	♀	XX		589				
5,148	♀	XX						1261
5,212	♀	XX						679
5,089	♀	XX						917
4,985	♀	XX						1026
5,211	♀	XX						773
594	♀	X*Y			145*			
599	♀	X*Y	223					
5,038	♀	X*Y					547	
4,994	♀	X*Y					340	
5,020	♀	X*Y					455	
5,091	♀	X*Y					295	
5,019	♀	X*Y					444	
592	♀	X*X		733				
596	♀	X*X		597		813		
595	♀	XO	245		397			
591	♀	X*O	235					
587	♀	X*XY				834		
589	♂†	X*XY				528		
Mean ± s.d.			235 ± 5	621 ± 38	376‡	751 ± 74	407 ± 26	931 ± 102
Ratio 1X:2X				1/2.65		1/2.00		1/2.28
Mean protein concentration (µg per 50 µl of homogenate)			146	171	158‡	180	211	291

STS activity is expressed as pmol dehydroepiandrosterone sulphate cleaved per mg protein per h. For determination of STS, early confluent cells were collected with trypsin, washed twice in Hank's balanced salt solution, suspended in lysis buffer (5 mM Tris-HCl, 2 mM EDTA, 7 mM mercaptoethanol, 0.1% Triton X-100 pH 7.5) and sonicated using a Branson sonifier with microtip (4 × 5 s, 50 W). ³H-dehydroepiandrosterone sulphate (DHEAS, specific activity 24 mCi µmol⁻¹; NEN) was diluted with cold DHEAS to a specific activity of 10 µCi µmol⁻¹, then extracted three times with diethyl ether to remove free dehydroepiandrosterone. 10 µl of this solution, corresponding to 10 nmol DHEAS, were incubated at 37 °C for 2 h with 50 µl of homogenate and 50 µl of 0.1 M Tris-HCl buffer pH 7.5, containing 10 mg bovine serum albumin (BSA) per ml. Then 0.1 ml of 0.2 M NaOH and a 10-fold excess of cold DHEAS were added, followed by extraction with 3 vol of ether. An aliquot of the ether phase was evaporated to dryness, and radioactivity determined by scintillation counting²². Blank controls contained lysis buffer with 10 mg ml⁻¹ BSA instead of homogenate. The relationship between STS activity and protein concentration is linear ($r = 0.99 \pm 0.07$) up to 800 µg of protein. Mean enzyme levels can vary in subsequent tests due to non-saturating substrate concentrations in the assay. Each value represents the mean of two independent cultures which were each tested in duplicate.

* Cultures not confluent. † Sterile male. ‡ Confluent cultures only.

X* chromosome, strongly argues against a duplication or deletion of the STS locus on the abnormal X*. It follows that, in this species, either the STS gene is not closely linked to gene(s) responsible for sex reversal, or the two genes have been separated by rearrangement on the aberrant X* chromosome.

We thank Professor Karl Fredga for providing material for this study; A. Gullberg, K. Jönsson, B. Müller-Migl, U. Strohmaier, M. Sunner and J. Zimmer for technical help; and Professor Ulrich Wolf for reading the manuscript. This work was

supported by grants from the Deutsche Forschungsgemeinschaft, the Swedish Natural Science Research Council and the Deutsche Akademischer Austauschdienst.

Received 25 November 1981; accepted 5 March 1982.

Table 2 G6PD activity (mU per mg protein) in wood lemming cultures with XY, X*Y and XX sex chromosome constitutions

	XY		X*Y		XX
Animal no.		Animal no.		Animal no.	
4,360	44.5	4,994	33	4,985	47
4,988	38.5	5,031	47	5,148	54
4,992	47	5,020	44	5,211	31
4,656	44	5,091	49	5,089	44
4,849	49	5,019	47	5,212	57
Mean±s.d.	44.6±1.8		44.0±2.9		46.6±4.5
		44.3±1.6			46.6±4.5

G6PD was assayed according to Löhr and Waller²³. Each value represents the mean of two independent cultures which were each tested twice.

1. Lyon, M. F. *Nature* **190**, 372-374 (1961).
2. Mohandas, T., Shapiro, L. J., Sparkes, R. S. & Sparkes, M. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5779-5783 (1979).
3. Tiepolo, L. *et al. Hum. Genet.* **54**, 205-206 (1980).
4. Wolf, U. *et al. Hum. Genet.* **54**, 149-154 (1980).
5. Wachtel, S. S., Koo, G. C., Breg, W. R. & Genel, M. *Hum. Genet.* **56**, 183-197 (1980).
6. Ohno, S. *Sex Chromosomes and Sex-linked Genes* (Springer, Berlin, 1967).
7. Müller, C. R., Migl, B., Traupe, H. & Ropers, H. H. *Hum. Genet.* **54**, 197-199 (1980).
8. Lifschytz, E. & Lindley, D. L. *Proc. natn. Acad. Sci. U.S.A.* **69**, 182-196 (1972).
9. *Rep. of Committee on Genetic Constitution of the X and Y, Human Gene Mapping 6*, 1981 (The National Foundation, New York, in the press).
10. Ferguson-Smith, M. Presented at the 6th *Hum. Gene Mapping Conf.*, Oslo 1981.
11. Bedin, M., Weil, D., Fournier, T., Cedard, L. & Frezal, J. *Hum. Genet.* **59**, 256-258 (1981).
12. Ropers, H. H. *et al. Hum. Genet.* **57**, 354-356 (1981).
13. Herbst, E. W., Fredga, K., Frank, F., Winking, H. & Gropp, A. *Chromosoma* **69**, 185-191 (1978).
14. Fredga, K., Gropp, A., Winking, H. & Frank, F. *Nature* **261**, 225-227 (1976).
15. Gropp, A., Winking, H., Frank, F., Noack, G. & Fredga, K. *Cytogenet. Cell Genet.* **17**, 343-358 (1976).
16. Gropp, A., Fredga, K., Winking, H. & Frank, F. *Ann. Biol. anim. Biochem. Biophys.* **18**(2), 367-375 (1978).
17. Ropers, H. H., Migl, B., Zimmer, J. & Müller, C. R. *Cytogenet. Cell Genet.* **30**, 168-173 (1981).
18. Polani, P. E. *Hum. Genet.* (in the press).
19. Ferguson-Smith, M. A. *J. med. Genet.* **2**, 142 (1965).
20. Lyon, M. F. *Proc. R. Soc. B187*, 243-268 (1974).
21. Goodfellow, P. N. & Tippet, P. *Nature* **289**, 404-405 (1981).
22. Ropers, H. H. *et al. Hum. Genet.* **57**, 354-356 (1981).
23. Löhr, G. W. & Waller, H. G. in *Methoden der Enzymatischen Analyse* (ed. Bergmeyer, H. U.) 675 (Chemie, Weinheim, 1974).

Transposition of a tandem duplication of yeast mating-type genes

James E. Haber & David T. Rogers

Rosenstiel Basic Medical Sciences Research Center and Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA

The switching of mating-type genes in both homothallic and heterothallic strains of the yeast *Saccharomyces cerevisiae* involves a site-specific transposition event¹⁻³. A sequence at the mating-type locus (*MAT*) is replaced by a copy of *a* or α information from the unexpressed loci, *HML* or *HMR*. Although *MATa* and *MAT α* contain unique sequences of ~700 base pairs (bp), they are flanked by sequences that are also found at *HML* and *HMR*^{4,5} (Fig. 1a). Recent results⁶⁻⁸ have suggested that the switching of *MAT* alleles involves an intrachromosomal mitotic gene conversion event, in which a donor sequence at *HML* or *HMR* pairs with the homologous sequences at the *MAT* locus. On this view the switching of mating-type information must be able to accommodate the region of non-homology in a heteroduplex structure. (Analogous gene conversions of non-homologous regions have been well documented during meiosis at *MAT* and for deletions in other yeast genes^{9,10}.) One prediction of a gene conversion mechanism is that it should be possible to transpose a tandem duplication of mating-type genes from *HMR* to *MAT*. In such a switch, pairing would occur between *MAT* and homologous sequences flanking the two copies of the tandem duplication at *HMR*. Our results reported here demonstrate that a tandem duplication created at *HMR* by recombinant DNA techniques can be transposed to *MAT* without the loss of these sequences at the donor locus.

Tandem duplications of mating-type genes at *HMR* were constructed by transforming the heterothallic (*ho*) strain DB 745 (*HML α MAT α HMRa ura3-52 leu2 ade1*) with plasmid pDRMATa3, which contains the 3.5 kilobase (kb) *EcoRI*-*HindIII* *MATa* fragment and the 1.1 kb *HindIII* *URA3* fragment inserted into pBR322 (Fig. 1a). In 3 of 20 Ura⁺ transformants the plasmid had integrated at *HMR*. These were shown by Southern blot analysis to be two types. In DR103 and DR107 a single copy of the plasmid was inserted, generating the structure *HMR::(MATa-URA3-pBR322)* (Fig. 1b). At least one of the two fusions between *MAT* and *HMR* sequences expressed *a* information, so that the *MATa* cell was phenotypically non-mating. The third transformant had two copies of the plasmid tandemly inserted, generating the structure *HMR::(MATa-URA3-pBR322)₂* (Fig. 1c). Again, at least one of the *a*-containing regions expressed *a* information; thus strain DR204 was also non-mating. We have also obtained an insertion of pDRMATa3 at *MATa*, giving the structure *MATa-URA3-pBR322-MATa* (strain DR212, Fig. 1d). This duplication has the structure expected for the transposition of a tandem duplication for *HMR::(MATa-URA3-pBR322)*.

Figure 2a shows an example of the Southern blot analysis used to confirm these structures. A *Bam*HI digest of DNA from the transformants was probed with labelled pDRMATa3 (Fig. 2a, lanes 1-4) or with labelled pBR322 (lanes 5-8). Because there is only a single *Bam*HI site in pDRMATa3 (Fig. 1), one would expect a single integrant of the plasmid to yield two *Bam*HI fragments, one of which should hybridize strongly with pBR322, and the other, homologous to only 300 bp (Fig. 1b), weakly. This was indeed the case for DNA from transformants DR103 and DR107 (Fig. 2a, lanes 6, 7). The third transformant, DR204, gave rise to the same *Bam*HI fragments plus one strongly hybridizing fragment the same size as the original plasmid (Fig. 2a, lane 8). This is consistent with the *HMRa::(MATa-URA3-pBR322)₂* structure shown in Fig. 1c. The tandem duplication at *MAT* (strain DR212) results in differently sized *Bam*HI fragments (Fig. 2a, lane 1).

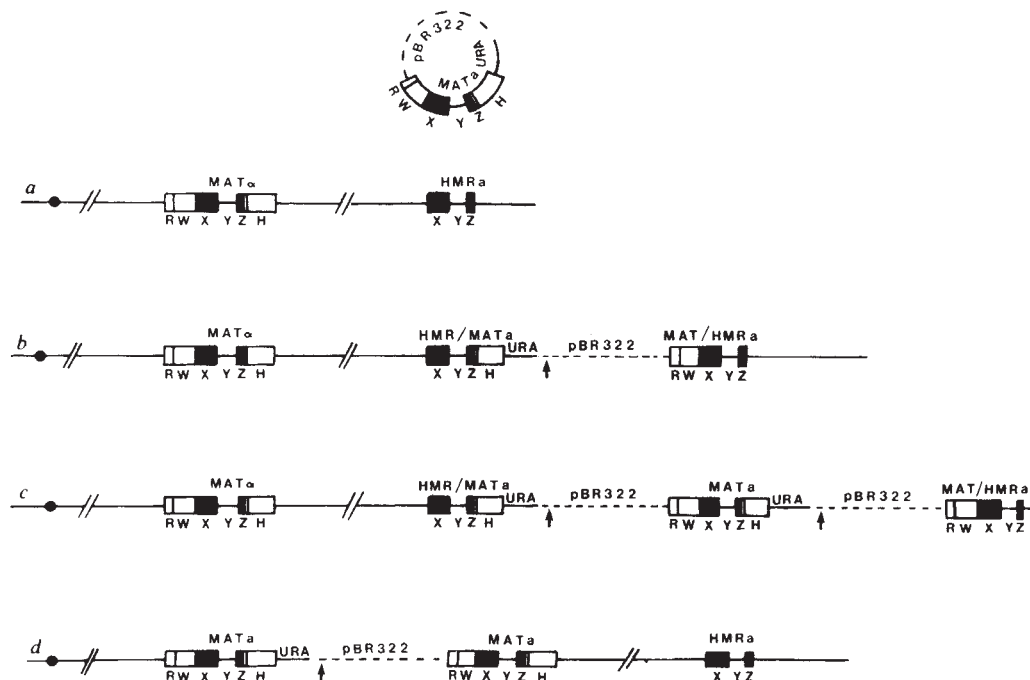
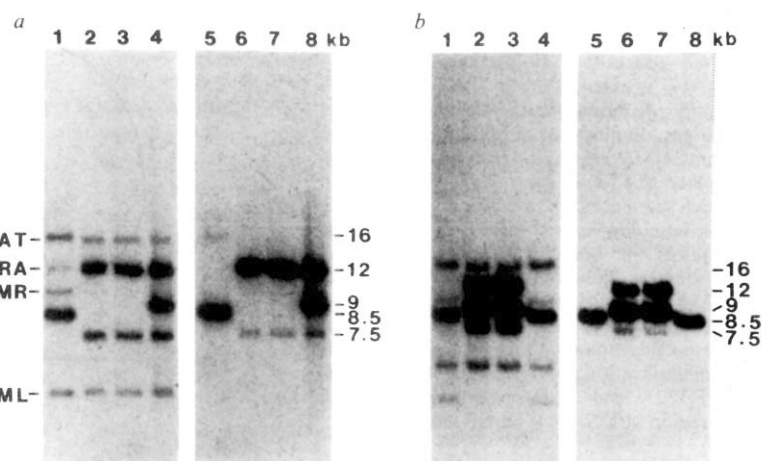


Fig. 1 Structures generated by the integration of plasmid pDRMATa3 by homologous recombination at *HMRa* or *MATa*. The plasmid contains the 3.5 kb *EcoRI*-*HindIII* fragment including *MATa* and the 1.1 kb *HindIII* fragment of *URA3*. *a*, The regions of homology between the plasmids and *HMRa* are indicated as solid blocks (X, Y and Z, according to ref. 5). The regions of homology between the plasmid and *MATa* are shown by the open and solid blocks [the W region plus the adjacent segments extending to the *EcoRI* site (R) and the *HindIII* site (H)]. The *HMLa* locus on the left arm of the chromosome (not shown) shares with *MAT* the homology regions W, X and Z. The *Ya* region at *MATa* is not homologous to *Ya* in the plasmid or at *HMRa*. The pBR322 sequences are indicated as ---. *Bam*HI sites in pBR322 used to analyse the structure of the transformants are shown by arrows. *b*, A single integration of the plasmid at *HMRa* (strains DR103 and DR107). *c*, A tandem insertion of the plasmid at *HMRa* (strain DR204). *d*, Insertion of the plasmid at *MATa* (strain DR212).

Fig. 2 Southern blot analysis of a transposition of a tandem duplication of *a*-mating-type sequences from *HMRa::(MATa-URA3)₂* to replace *MATa*. *a*, Analysis of insertions of pDRMATA3 at *HMRa* and *MATa*. DNAs from four strains, described below, were digested with *Bam*HI, separated by agarose gel electrophoresis and the Southern blot probed with either labelled pDRMATA3 (lanes 1–4) or labelled pBR322 (lanes 5–8). Lanes 1, 5: strain DR212 (Fig. 1*d*), *MATa::(MATa-URA3)*; lanes 2, 6: strain DR103 (Fig. 1*b*), *HMRa::(MATa-URA3)*; lanes 3, 7: strain DR107 (Fig. 1*b*), *HMRa::(MATa-URA3)*; lanes 4, 8: strain DR204 (Fig. 1*c*), *HMRa::(MATa-URA3)₂*. The identity of each band was determined by additional hybridizations with DNA probes specific for *URA3*, *MAT* and *HML* (data not shown). Two *Bam*HI fragments of 12 kb and 7.5 kb are generated when pDRMATA3 integrates at *HMRa* (see Fig. 1*b* and lanes 2, 3, 6, 7). A tandem insertion of the plasmid at *HMRa* creates the same two *Bam*HI fragments, plus a 9 kb fragment identical in size to the plasmid (see Fig. 1*c* and lanes 4, 8). Creation of a tandem duplication at *MATa* (by the integration of plasmid pDRMATA3 at *MATa*) creates two *Bam*HI fragments of 16 and 8.5 kb, both of which hybridize to labelled pBR322. *b*, Southern blot analysis of four meiotic segregants from a diploid (DR204/A145 no. 2) of genotype:

HMLa MATa URA3 pBR322 MATa HMRa::(MATa-URA3)₂
HMLa MATa HMRa

DNA from each of four segregants was cut with *Bam*HI and the separated fragments probed with labelled pBR322 (lanes 1–4) or pDRMATA3 (lanes 5–8). Two of the segregants were *a*-mating and *Ura*⁺, and carry *MATa* sequences derived by transposition from *HMRa::(MATa-URA3)₂*. These two strains (lanes 1, 4; lanes 5, 8) contain *Bam*HI fragments characteristic of a tandem duplication at *MATa* (see Fig. 2*a*, lanes 4, 8). The other two segregants (lanes 2, 3; lanes 6, 7) were non-mating, *Ura*⁺ and carry *MATa* as well as *a* information expressed at *HMRa::(MATa-URA3)₂*. The *Bam*HI fragments are identical in size to those in the original parent strain, DR204 (Fig. 2*a*, lanes 3, 7).



Transposition of copies of *a* information from *HMRa::(MATa-URA3-pBR322)* was monitored by crossing the non-mating *Ura*⁺ strains with strain A145 (*ho MATa ura3 leu1 ade1*) and selecting *Ura*⁺ diploids. These could only arise if the non-mating strains became *a*-mating by the transposition of *a* information from the tandem duplication at *HMR* to replace *MATa*. Ten diploids from each of the three mating experiments were analysed by tetrad analysis to determine whether the conversion of the *MATa HMRa::(MATa-URA3-pBR322)* strain to *MATa* resulted from an event that also transposed a copy of the *URA3* gene (see Table 1). In all 20 cases involving strains DR103 and DR107 the segregation of markers indicated that a single copy of *MATa* had been transposed to the *MAT* locus without *URA3*—that is, the expression of an *a* mating type was not linked to *URA3* and each tetrad contained two *Ura*⁺ and two *Ura*[−] segregants. If there were two unlinked *URA3* genes segregating in these cases, the probability of finding a 2⁺:2[−] tetrad would be only 1/6. Furthermore, there were an equal number of *a*-mating *Ura*[−] strains and non-mating *Ura*⁺ segregants, presumably carrying *MATa* and either *HMRa* or *HMRa::(MATa-URA3-pBR322)* respectively. In contrast, in two of the ten diploids isolated from mating strain DR204 with A145 there was a second copy of *URA3* segregating in the cross and one copy of *URA3* was tightly linked to *MATa*. This is evident from the fact that most tetrads contained three or four *Ura*⁺ segregants per tetrad and that every *a*-mating segregant was *Ura*⁺. Half of the *MATa* segregants were *Ura*⁺ and non-mating. We concluded that in two of the ten transpositions of *a* information from *HMRa::(MATa-URA3-pBR322)₂* in strain DR204 the transposition event also mobilized the *URA3* gene.

Using Southern blot analysis to study one transposition event (DR204/A145 no. 2), we have shown that (1) the transposition did indeed create a tandem duplication of *MATa* flanking *URA3* and pBR322 at *MAT*; and (2) that the structure at the donor *HMRa::(MATa-URA3-pBR322)₂* locus was not altered (see Fig. 2*b*). DNA was isolated from one tetrad of diploid DR204/A145 no. 2 in which all four segregants were *Ura*⁺. A Southern blot of DNA digested with the restriction endonuclease *Bam*HI was probed with labelled pDRMATA3 (Fig. 2*b*, lanes 1–4) or pBR322 (lanes 5–8). The two segregants that were *a*-mating and *Ura*⁺ contained two new bands (16 and 8.5 kb) characteristic of an insertion of the *MATa-URA3-pBR322* plasmid at *MATa* (lanes 1, 4). The other two

segregants which were non-mating and *Ura*⁺ were shown to carry *MATa* and the pair of 7.5 and 12 kb *Bam*HI restriction fragments found with the *HMRa::(MATa-URA3-pBR322)₂* duplication of the parent strain DR204 (lanes 2, 3). These data confirm that the transposition event created a tandem duplication of *MATa* and transposed the intervening *URA3* and pBR322 sequences.

These experiments show that the mechanism of switching *MAT* genes can transpose a tandem duplication of mating-type genes and 5.5 kb of non-homologous intervening sequences. If the evidence favouring a gene conversion model for *MAT* switching is correct, even the transposition of a single copy of *a* to replace *MATa* must form a heteroduplex in which the non-homologous sequences of *a*- and *α*-specific DNA are not paired. In the case where a tandem duplication of *MAT* genes was transposed, we presume that a heteroduplex formed between homologous sequences (W and Z) on the outside

Table 1 Heterothallic conversions of *MATa* to *MATa* or to *MATa-URA3-pBR322-MATa*

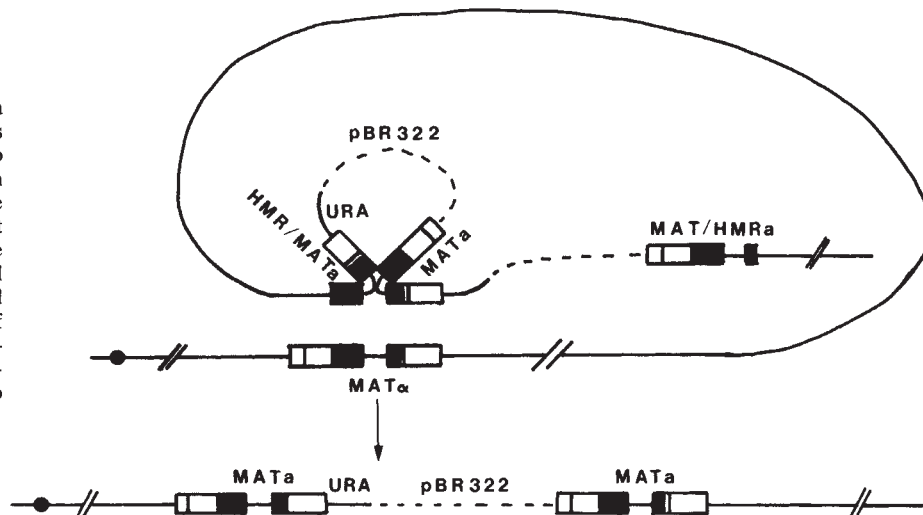
	Tetrad type				
	Mat Ura	Mat Ura	Mat Ura	Mat Ura	Mat Ura
	<i>a</i> −	<i>a</i> +	<i>a</i> +	<i>a</i> +	<i>a</i> +
	<i>a</i> −	<i>a</i> −	<i>a</i> +	<i>a</i> +	<i>a</i> +
	NM +	<i>a</i> −	<i>a</i> −	<i>a</i> −	NM +
Diploid	NM +	NM +	<i>a</i> +	NM +	NM +
DR103/A145*	8	36	13		
DR107/A145*	7	37	10		
DR204/A145†	5	26	8		
No. 2			1	8	1
No. 8			3	11	2

Heterothallic non-mating strains DR1203 and DR107 (*ho, MATa, HMRa::(MATa-URA3-pBR322)*) and strain DR204 (*ho, MATa, HMRa::(MATa-URA3-pBR322)₂*) were crossed with the *MATa HMRa ura3* strain A145. Ten independent rare matings of each cross were purified, sporulated and analysed by tetrad dissection. The total numbers of each type of tetrad are shown. Non-mating segregants are designated NM.

* Sum of 10 independent diploids analysed.

† Sum of eight independent diploids analysed.

Fig. 3 A model for the transposition of a tandem duplication of mating type genes from *HMR::(MATa-URA3-pBR322)₂* to replace *MATα*. A gene conversion event in which sequences at *HMR* are used to replace non-homologous sequences at *MAT* first requires the homologous pairing of these regions, as shown. Instead of the normal 0.7 kb region of non-homology transferred between *HMRa* and *MATα*, the pairing of the outside homologous regions of the tandem duplication at *HMRa::(MATa-URA3-pBR322)₂* leads to the insertion of a 7.9 kb region of DNA not homologous to *Yα*.



portions of two different donor sequences (Fig. 3) could then create a large non-homologous region in the middle of the heteroduplex. DNA mismatch repair favouring the incoming strand would then allow the entire region to be transposed. It is significant that we found such transpositions with strain DR204 but not with strain DR103 or DR107. One explanation for this difference is that the extent of homologous pairing that can be formed, especially on the right side of the unique sequences, is substantially greater (1.1 kb compared with 0.23 kb) using the multiple tandem duplication *HMRa::(MATa-URA3-pBR322)₂* than the single *HMRa::(MATa-URA3-pBR322)* integrant. The greater homology may stabilize the unusually large non-homologous region in a heteroduplex structure.

We have recently also carried out a complementary study of the switching of a tandem duplication of *MAT* genes at the mating-type locus (J.E.H. and D.T.R., in preparation). Homothallic strains carrying a *MATa* duplication switched efficiently and usually the tandem duplication was replaced by a single *MATa* locus. Pedigree studies have shown that the loss of the duplication and the switch to the opposite mating type occurs during a single cell division. Analogous to our interpretation of transposition of a duplication from *HMR* to *MAT* (Fig. 3), we explain the complementary study by supposing that the donor *HMRa* must pair with the common region X, to the left of one *MATα* gene and the common region Z1 on the right of the other *MATα* locus. A gene conversion would then replace a large (7.9 kb) non-homologous sequence with the 0.7 kb unique *a* sequence.

We believe our results are consistent with an intrachromosomal gene conversion mechanism for switching mating type.

This research was supported by USPHS grant GM20056. We thank Barbara Weiffenbach for her comments on the manuscript. D.T.R. was a Charles A. King Trust Research Fellow.

Induction of a DNA glycosylase for *N*-methylated purines is part of the adaptive response to alkylating agents

Peter Karran*, Thomas Hjelmgren†
& Tomas Lindahl*

* Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, London NW7 1AD, UK

† Department of Medical Biochemistry, Gothenburg University, Gothenburg, Sweden

The resistance of *Escherichia coli* to simple alkylating agents, for example *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG), is markedly increased in cells previously exposed to low doses of the agents. This adaptive response seems to reflect improved repair of alkylation lesions in DNA, and cells become less sensitive to both the mutagenic and killing effects of alkylating agents^{1,2}. Adaptation to the former is due largely to the induction of a repair function that removes *O*⁶-methylguanine (*m*⁶G) from DNA^{3,4}. This activity has been purified from induced cells and is attributed to a protein that transfers the methyl group of an *m*⁶G residue in DNA to one of its own cysteine residues⁵. Consequently, no excision of damaged nucleotide residues seems to be necessary. The adaptive response to killing by alkylating agents, however, is less understood. DNA excision-repair with gap-filling catalysed by DNA polymerase I seems to occur as part of the inducible response, as it depends on the presence of a functional *polA*⁺ gene² and enhanced DNA polymerase I-mediated repair synthesis has been observed in adapted cells challenged with MNNG⁶. Here we show that the adaptive response involves induction of not only a methyltransferase for *m*⁶G, but also a DNA glycosylase that catalyses the release of purines methylated at ring nitrogen atoms. We also identify the minor alkylation product 3-methylguanine (*m*³G) as an important lethal lesion.

One of the major methylation products in DNA, 3-methyladenine (*m*³A), is known to be actively and rapidly removed *in vivo*^{7,8}. The enzyme responsible, 3-methyladenine DNA glycosylase I, can easily be detected in extracts from unadapted cells, and it seems to be constitutively expressed. This activity has been extensively purified and characterized⁹. The enzyme, which has a molecular weight of about 19,000, shows a very narrow substrate specificity and of several different DNA methylation products, only *m*³A is released from methylated DNA. *E. coli* mutants (*tagA*) deficient in this enzyme have been isolated, and although they are less able to remove *m*³A from their DNA after exposure to alkylating agents⁸, they are not totally deficient in *m*³A DNA glycosylase activity. The

Received 14 September 1981; accepted 23 February 1982.

- Oshima, Y. & Takano, I. *Genetics* **67**, 327-335 (1971).
- Hicks, J. B., Strathern, J. N. & Herskowitz, I. *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A., Shapiro, J. & Adhya, S.) 457-462 (Cold Spring Harbor Laboratory, New York, 1977).
- Hicks, J., Strathern, J. & Klar, A. J. S. *Nature* **282**, 478-483 (1979).
- Nasmyth, K. & Tatchell, K. *Cell* **19**, 753-764 (1980).
- Strathern, J. N., Spatola, E., McGill, C. & Hicks, J. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2839-2843 (1980).
- Haber, J. E., Rogers, D. T. & McCusker, J. H. *Cell* **22**, 277-289 (1980).
- Klar, A. J. S., McIndoo, J., Strathern, J. N. & Hicks, J. B. *Cell* **22**, 291-298 (1980).
- Haber, J. E., Rowe, L. & Rogers, D. *Molec. cell. Biol.* **1**, 1106-1119 (1981).
- Klar, A. J. S., Fogel, S. & Lusnak, K. *Genetics* **92**, 777-782 (1979).
- Fink, G. R. & Styles, C. A. *Genetics* **77**, 231-244 (1974).

Table 1 DNA glycosylase activities in various *E. coli* strains before and after adaptation by MNNG treatment

<i>E. coli</i> strain	DNA Glycosylase	Specific activity of cell extract ($\mu\text{U mg}^{-1}$)	
		Unadapted	Adapted
<i>E. coli</i> B/r	Uracil	3,800	3,500
	Hypoxanthine	2.1	2.0
	3-Methyladenine I	3.6	3.6
	3-Methyladenine II	0.22	4.1
	3-Methylguanine	~0.05	1.1
	7-Methylguanine	~0.02	0.2
	Formamidopyrimidine	4.2	3.9
	Thymine glycol	1.9	1.7
	Urea	5.5	4.8
<i>E. coli</i> BK2114 (<i>tagA</i>)	3-Methyladenine I	<0.02	<0.1
	3-Methyladenine II	0.20	7.8
	3-Methylguanine	~0.05	1.7
	7-Methylguanine	~0.02	0.4

E. coli cells were adapted by growth in MNNG ($1 \mu\text{g ml}^{-1}$) for 60 min. The uracil, hypoxanthine, formamidopyrimidine, thymine glycol and urea DNA glycosylases were assayed in cell-free extracts as described previously^{13,22-25}. The products released by $m^3\text{G}$, $m^7\text{G}$ and $m^3\text{A}$ DNA glycosylases from ^3H -dimethylsulphate-treated DNA were characterized by HPLC as described in Fig. 2 legend.

residual function, which represents 5–10% of the total activity in wild-type cell extracts, has properties different from $m^3\text{A}$ DNA glycosylase I: it is resistant to product inhibition by free $m^3\text{A}$; it is relatively heat resistant; and it is present in similar amounts in wild-type cells and in *tagA* mutants. It has therefore been concluded that the minor activity is a separate enzyme, which has been termed 3-methyladenine DNA glycosylase II^{8,10}.

When the enzymatic release of free $m^3\text{A}$ from alkylated DNA was measured, using cell extracts from unadapted and adapted *E. coli*, the adapted cells exhibited two- to threefold more DNA glycosylase activity. To assess the relative contributions of $m^3\text{A}$ DNA glycosylases I and II, the cell extracts were chromatographed on Sephadex G-75, and column fractions assayed in the presence and absence of $m^3\text{A}$ (3 mM). In these conditions, $m^3\text{A}$ DNA glycosylase I is 90–95% inhibited whereas enzyme II is not affected. Figure 1a shows that enzyme II accounted for 5–10% of the total enzyme activity in extracts of unadapted cells, in agreement with previous results⁸. In adapted cells, however, this enzyme was responsible for 50–70% of the total activity (Fig. 1b). Thus, although there was no detectable difference between the levels of $m^3\text{A}$ DNA glycosylase I in adapted and unadapted cells, $m^3\text{A}$ DNA glycosylase II was present at a 20-fold higher level in adapted cells. This induced activity was eluted before the constitutively expressed enzyme I on Sephadex G-75 (Fig. 1b), indicating a higher molecular weight of enzyme II.

The identity of the induced enzyme as $m^3\text{A}$ DNA glycosylase II was confirmed by experiments with an *E. coli tagA* mutant (BK2114), which lacks enzyme I. In this strain, the level of $m^3\text{A}$ DNA glycosylase activity was low, but could be induced 20- to 40-fold by exposure to MNNG (Table 1). The induced activity was resistant to inhibition by free $m^3\text{A}$, as expected for enzyme II, and also exhibited the heat stability of enzyme II. The separate identities of enzyme I and the inducible enzyme II is further supported by the observation of Evensen and Seeberg that *E. coli alk* mutants are deficient specifically in the latter enzyme¹¹. In agreement with previous studies of the adaptive response², MNNG was a good inducer of $m^3\text{A}$ DNA glycosylase II, whereas methyl methanesulphonate was less efficient and 4-nitroquinoline-1-oxide had no effect. An *E. coli* strain (BS21) which expresses the adaptive response constitutively and contains high levels of the methyltransferase acting on $m^6\text{G}$ (ref. 12), did not contain a similarly elevated amount of $m^3\text{A}$ DNA glycosylase II, although it had slightly more (two- to threefold) enzyme than did unadapted wild-type cells.

In addition to the increased release of $m^3\text{A}$ from alkylated DNA, extracts from adapted cells were also active on other *N*-methylated purines in DNA. Thus, whereas only extremely low levels of DNA glycosylase(s) which released $m^3\text{G}$ and 7-methylguanine ($m^7\text{G}$) from methylated DNA were detectable in extracts of unadapted cells, such activities were easily demonstrable in crude extracts of adapted cells (Fig. 2, Table 1). Even in the latter extracts, however, $m^7\text{G}$ DNA glycosylase activity was quite low; in assay conditions which resulted in complete release of $m^3\text{A}$ and $m^3\text{G}$ from methylated DNA, only ~0.3% of the $m^7\text{G}$ was enzymatically released.

Several other DNA glycosylases, which specifically recognize different types of damaged bases in DNA, have been found in *E. coli*. A survey of the levels of these activities in extracts from unadapted and adapted cells showed that only $m^3\text{A}$ DNA glycosylase II, $m^3\text{G}$ DNA glycosylase and $m^7\text{G}$ DNA glycosylase activities were induced during adaptation by MNNG treatment (Table 1). Note that the amount of formamidopyrimidine DNA glycosylase was not increased, although this enzyme acts on a major secondary alkylation product¹³.

The inducible DNA glycosylase activities for $m^3\text{A}$, $m^3\text{G}$ and $m^7\text{G}$ appear to be associated with a single enzyme as the relative levels of induction of the three activities were the same (10- to 40-fold) in different experiments, although the small amounts of $m^3\text{G}$ and $m^7\text{G}$ DNA glycosylase activity in unadapted cells made accurate measurements difficult. (Neither activity was as efficiently induced as the methyltransferase for $m^6\text{G}$ [ref. 14].) Furthermore, we have been unable to separate the inducible DNA glycosylase activities by a variety of chromatographic procedures (DEAE-cellulose, Sephadex gel chromatography, DNA-cellulose).

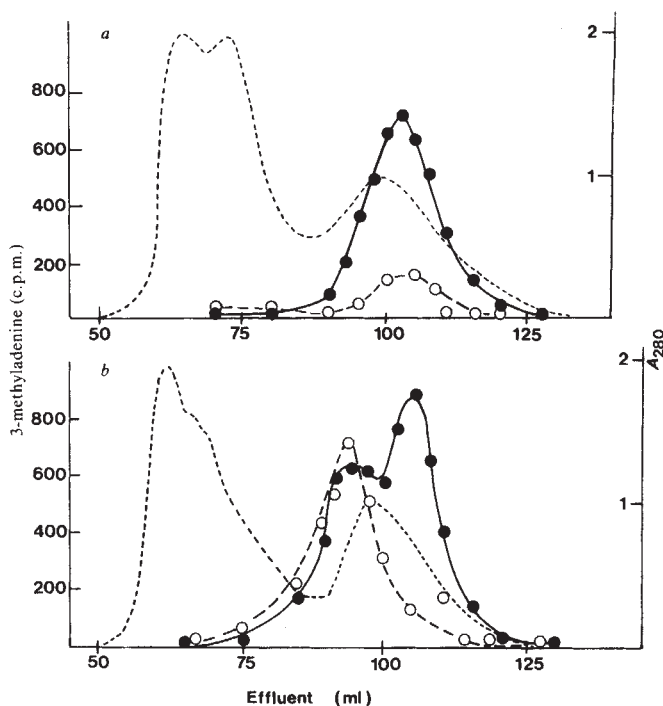


Fig. 1 3-Methyladenine DNA glycosylases I and II in *E. coli* B/r before and after exposure to MNNG. Cells were adapted by growth in MNNG ($1 \mu\text{g ml}^{-1}$) for 60 min. Extracts from 3 g cells were prepared by grinding with sand and partially purified by streptomycin and ammonium sulphate fractionation⁹. Partially purified extracts were chromatographed on Sephadex G-75 (1 cm diameter $\times$ 100 cm) in 0.3 M NaCl, 0.05 M Tris-HCl pH 8.0, 1 mM EDTA, 1 mM 2-mercaptoethanol, 5% glycerol. Duplicate 10- μl aliquots of each fraction were assayed in 100 μl standard assay mixture⁹ containing ^3H -dimethylsulphate-treated DNA (5,000 c.p.m.), with or without product inhibition of $m^3\text{A}$ DNA glycosylase I by free $m^3\text{A}$ (3 mM). *a*, Unadapted cells; *b*, adapted cells. ---, A_{280} ; ●, assay without free $m^3\text{A}$; ○, assay in the presence of 3 mM $m^3\text{A}$.

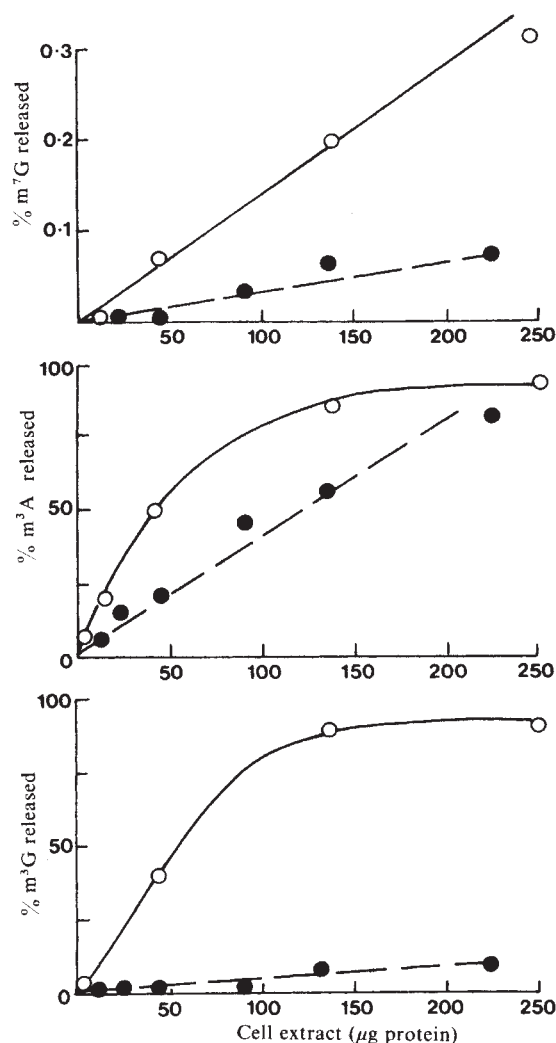


Fig. 2 DNA glycosylase activities for methylated purines in extracts of unadapted and adapted *E. coli*. Cultures (10 ml) of *E. coli* B/r were adapted by growth in MNNG ($1 \mu\text{g ml}^{-1}$) for 60 min. Cells were collected by centrifugation, washed and disrupted by sonication in 0.07 M Tris-HCl pH 8.0, 1 mM EDTA, 1 mM dithiothreitol, 5% glycerol. Debris were removed by centrifugation (10,000 g, 5 min) and the crude extract used to assay the release of methylated purines. The assay mixture (100 μl) contained: 0.07 M HEPES-KOH pH 7.8, 1 mM EDTA, 1 mM dithiothreitol, 5% glycerol, 2 μg ^3H -dimethylsulphate-treated *M. luteus* DNA (50,000 c.p.m.) and 0–250 μg extract protein. After 20 min at 37 °C, mixtures were cooled to 0 °C, precipitated with 2 volumes of ethanol and centrifuged (10 min, 10,000 g). Supernatants were concentrated by evaporation to $\sim 30 \mu\text{l}$, supplemented with authentic methylated purines and analysed by HPLC using a Varian model 5000 chromatograph. Separation was on a Whatman Partisil 10/25 SCX column using a gradient of ammonium formate, pH 4.0, in 8% methanol essentially as described in ref. 21. The identities of released methylpurines were confirmed by paper chromatography on Whatman 3MM paper. Chromatograms were developed in methanol/ethanol/HCl/H₂O (50:25:6:19) for 24 h. Radioactivity associated with the positions of authentic marker compounds was determined as described previously⁸. ●, Unadapted cells; ○, adapted cells.

Table 2 Excision of methylated purines from the DNA of dimethylsulphate-treated *E. coli*

Incubation time after alkylation (min)		Alkylated bases remaining in DNA (pmol per 10 ¹¹ cells)				
		m ⁷ G	m ³ A*	m ³ G	m ⁷ A	m ¹ A
Unadapted	{ 0	124.7	0.93	1.61	4.15	0.34
	{ 30	106.7	0.92	0.78	2.55	0.18
Adapted	{ 0	114.5	0.94	0.68	3.20	0.38
	{ 30	109.1	0.54	<0.03	2.10	0.11

Cultures of *E. coli* B/r (1 l) in M9 medium containing glucose (0.4%) and casaminoacids (0.1%) were adapted as described in Table 1 legend. Unadapted cultures, grown in parallel, did not receive MNNG. Adapted and unadapted cells were collected by centrifugation, washed and resuspended in unsupplemented M9 (prewarmed to 37 °C) at a concentration of 4×10^{10} cells per ml. To both suspensions was added ^3H -dimethylsulphate (670 c.p.m. pmol⁻¹) to give a final concentration of 0.17 mM. After 10 min incubation at 37 °C, one-half of each suspension was diluted with 5 volumes of crushed ice and the cells collected by centrifugation. The remaining cells were diluted into full growth medium and incubated for a further 30 min before collection. DNA was extracted from cells as described in ref. 8, hydrolysed and alkylated purines analysed by HPLC as described in Fig. 2 legend.

* More than 95% of the m^3A initially formed was removed during the treatment with the alkylating agent⁸ and was not detected in the 0 min samples.

As the potentially lethal lesion m^3A is rapidly removed by the constitutively expressed enzyme I, it is unclear why an additional, inducible activity should be necessary. Further, the major alkylation lesion m^7G appears to be an innocuous form of damage, which usually persists for a long time in both unadapted and adapted cells, although the latter have slightly better capacity to remove this base (Fig. 2). Like m^3A , the minor DNA methylation product, m^3G (ref. 16), projects its methyl group into the narrow groove of the DNA double helix which, unlike the wide groove, is normally free of methyl groups. In contrast to m^3A , however, it is not efficiently removed in unadapted cells¹⁷. We have investigated the effect of adaptation on the rate of removal of this alkylation product *in vivo*. As shown in Table 2, m^3G was removed quite slowly in unadapted cells, with 50% remaining in DNA 30 min after treatment. In adapted cells, however, about one-half was removed immediately during the brief treatment with an alkylating agent, and by 30 min no residual m^3G was detectable. m^3A was excised with a similar high efficiency in unadapted and adapted cells, whereas most of the m^7G remained in the DNA in both cases. The minor alkylation products 7-methyladenine and 1-methyladenine, which were also resolved during HPLC of DNA hydrolysates, were inefficiently removed in adapted as well as unadapted cells (Table 2).

m^3G only accounts for about 1% of the total DNA lesions after treatment of cells with methylating agents, and is thus produced at a 15-fold lower level than m^3A . Since the latter, potentially lethal, lesion is rapidly removed by a constitutively expressed repair enzyme, similar amounts of m^3A and m^3G remain in DNA shortly after exposure. These lesions, both of which presumably can block the narrow groove of the double helix, would be expected to be of similar lethality. Consequently, the ability to repair m^3G in DNA would seem to be of great advantage for cells exposed to alkylating agents. As shown here, *E. coli* has a low level of such a repair activity, which can be augmented following exposure to alkylating agents. A m^3G DNA glycosylase activity has previously been observed in crude cell extracts of (unadapted) *Micrococcus luteus*¹⁸ and human lymphoblasts¹⁹. As m^3G is removed from DNA less efficiently than m^3A not only in unadapted *E. coli* but also in mammalian cells^{20,21}, it seems likely that m^3G will generally be responsible for a major part of the cytotoxic effect of methylating agents.

Support for this conclusion comes independently from Thomas *et al.*¹⁰, who partly purified enzyme II from unadapted *E. coli*. They found that the three DNA glycosylase activities were associated with a protein of molecular weight 27,000, and the m^3A - and m^7G -releasing activities were also shown to exhibit identical rates of heat inactivation, and to co-migrate on isoelectric focusing. The m^7G DNA glycosylase associated with enzyme II probably accounts for the low level of this activity previously reported in *E. coli*¹⁵.

We thank S. Stevens for valuable technical assistance at the MRC Cell Mutation Unit, University of Sussex, where part of this work was carried out, also Drs E. Seeberg and B. Sedgwick for bacterial strains, and Dr D. A. Goldthwait for a preprint of his paper. This work was supported in part by the Swedish MRC.

Received 23 December 1981; accepted 24 February 1982.

- Samson, L. & Cairns, J. *Nature* **267**, 281–283 (1977).
- Jeggo, P. *et al. Molec. gen. Genet.* **157**, 1–9 (1977); **162**, 299–305 (1977).
- Schendel, P. & Robins, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6017–6020 (1978).
- Karran, P., Lindahl, T. & Griffin, B. E. *Nature* **280**, 76–77 (1979).
- Olsson, M. & Lindahl, T. *J. biol. Chem.* **255**, 10569–10571 (1980).
- Billen, D. & Hellerman, G. R. *J. Bact.* **137**, 1439–1442 (1979).
- Lawley, P. D. & Orr, D. J. *Chem.-biol. Interact.* **2**, 154–157 (1970).
- Karran, P. *et al. J. molec. Biol.* **140**, 101–127 (1980).
- Riazuddin, S. & Lindahl, T. *Biochemistry* **17**, 2110–2118 (1978).
- Thomas, L., Yang, C. H. & Goldthwait, D. A. *Biochemistry* (in the press).
- Evensen, G. & Seeberg, E. *Nature* **296**, 773–775 (1982).
- Sedgwick, B. & Robins, P. *Molec. gen. Genet.* **180**, 85–90 (1980).
- Chetsanga, C. J. & Lindahl, T. *Nucleic Acids Res.* **6**, 3673–3684 (1979).
- Cairns, J. *Nature* **286**, 176–178 (1979).
- Laval, J., Pierre, J. & Laval, F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 852–855 (1981).
- Lawley, P. D., Orr, D. J. & Shah, S. A. *Chem.-biol. Interact.* **4**, 431–434 (1972).
- Lawley, P. D. & Warren, W. *Chem.-biol. Interact.* **12**, 211–220 (1976).
- Shackleton, J., Warren, W. & Roberts, J. J. *Eur. J. Biochem.* **97**, 425–433 (1979).
- Singer, B. & Brent, T. P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 856–860 (1981).
- Margison, G. P., Margison, J. M. & Montesano, R. *Biochem. J.* **157**, 627–634 (1976).
- Frei, J. V. *et al. Biochem. J.* **174**, 1031–1044 (1978).
- Lindahl, T. *et al. J. biol. Chem.* **252**, 3286–3294 (1977).
- Karran, P. & Lindahl, T. *J. biol. Chem.* **253**, 5877–5879 (1978).
- Demple, B. & Linn, S. *Nature* **287**, 203–208 (1980).
- Breimer, L. & Lindahl, T. *Nucleic Acids Res.* **8**, 6199–6211 (1980).

Adaptation to alkylation resistance involves the induction of a DNA glycosylase

Grethe Evensen* & Erling Seeberg†

* Institute of General Genetics, University of Oslo, Blindern, Oslo 3, Norway

† Division for Toxicology, Norwegian Defence Research Establishment, N-2007 Kjeller, Norway

Cells exposed to low doses of *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) or methyl methanesulphonate (MMS) acquire resistance to both the mutagenic and lethal effects of a challenging dose of the same agents (ref. 1, and ref. 2 for a review). This response, termed adaptation, has been ascribed to the induced synthesis and accumulation of *O*-6-methylguanine (m^6G) DNA methyltransferase^{3,4} which rapidly demethylates the m^6G residues induced by the challenging dose³⁻⁵. Mutant studies, however, indicate that mutagenic adaptation and killing adaptation are at least partly under different genetic control and may therefore involve the induction of different repair enzymes^{6,7}. Whereas mutagenic adaptation correlates with the induction of the transferase, the data presented here show that killing adaptation can be ascribed to the induction of a DNA glycosylase. This inducible glycosylase releases the alkylation product 3-methyladenine (m^3A) from DNA *in vitro* as does the constitutive m^3A DNA glycosylase previously characterized by Riazuddin and Lindahl⁸. However, the enzymes are encoded by different genes and appear to have different roles in DNA repair *in vivo*.

We have previously isolated from MNNG-mutagenized clones of *Escherichia coli* a mutant that is extremely sensitive to MMS, unable to reactivate alkylated phages and deficient in m^3A DNA glycosylase activity⁹. Further characterization of this strain (BK2106) has revealed that it carries two mutations, *tag* and *ada*, and not only lacks the glycosylase, but also is deficient in the induction of m^6G DNA methyltransferase. Elimination of the *ada* mutation in this strain by P1 transduction renders the cell wild-type resistant to the lethal effects of MMS in agar medium (Table 1). However, the *ada*⁺ *tag* transductant (BK2114) is still unable to reactivate alkylated phages (Fig. 1) and shows considerable sensitivity to MMS exposure in buffer (Fig. 2). These observations suggested that *E. coli* may have

two pathways for the repair of m^3A residues in DNA, one that is constitutively expressed and controlled by *tag*, and another that is inducible and controlled by *ada*. The inducible repair pathway would be turned on when the cells were exposed to alkylating agents in complete medium, but not when the cells were infected with alkylated phages. Table 1 shows that the adaptive treatment used to induce m^6G DNA methyltransferase also induces high levels of m^3A DNA glycosylase activity in the *tag* mutant (BK2114). Such induction is not observed in the *tag ada* double mutant (BK2106). Adapted BK2114 has also regained essentially normal ability to reactivate alkylated phages which implies that the inducible enzyme can replace the constitutive enzyme also in repair of phage DNA (Fig. 1). These results show that adaptation induces two different enzymes for repair of alkylated DNA, m^6G DNA methyltransferase and m^3A DNA glycosylase, and that both enzymes are under *ada* control.

Whereas constitutive m^3A DNA glycosylase (TagI) in wild-type cells is inhibited by m^3A (ref. 8), the inducible glycosylase (TagII) in BK2114 is unaffected by m^3A (Table 1). Adaptation of wild-type cells also results in the induction of m^3A DNA glycosylase activity not inhibited by m^3A , indicating that the inducible enzyme in BK2114 is not a mutated form of the normal *tag* gene product, but is indeed a separate enzyme. Karran *et al.*⁹ previously observed a minor m^3A -noninhibited m^3A DNA glycosylase in both wild-type cells and strain BK2106. This enzyme is more heat-stable than TagI and constitutes ~5% of the total m^3A DNA glycosylase activity in wild-type cells. It is probably the same as that which is induced to high levels in BK2114 by the adaptive treatment. In fact, Lindahl and co-workers¹⁰ have recently observed that the heat-stable activity of m^3A DNA glycosylase increases several-fold in wild-type cells during adaptation.

The observation that the *ada* mutation prevents induction of two different repair enzymes suggests that the *ada* gene is a control gene for the adaptive response and that structural genes for one or both of the enzymes may be found elsewhere on the chromosome. While searching for structural gene mutants we observed that the *alk-1* mutant^{11,12} was deficient in the induction of TagII, but proficient in induction of the transferase (Table 1). Similar results were obtained with another *alk* mutant (isolated by B. Duncan). No other enzyme defect has been linked to mutations in the *alk* gene. *alk* mutants are very sensitive to MMS exposure and it seems likely that the MMS sensitivity is caused by TagII deficiency, in which case the low levels of TagII activity present in wild-type cells must be essential for some repair not effected by TagI. It is possible

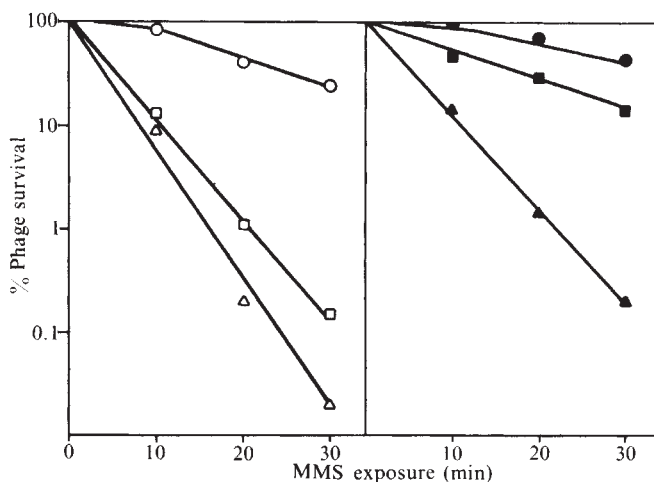


Fig. 1 Host cell reactivation of alkylated phages in non-treated and adapted cells. Phage λ cI857 were exposed to 0.05 M MMS for the times indicated and plated on exponentially growing untreated (open symbols) or adapted (closed symbols) cells of AB1157 (wild type, circles), BK2106 (*tag ada*, triangles) or BK2114 (*tag*, squares). Cells were adapted as described in Table 1 legend and conditions for the plating procedure were as published previously⁹.

† To whom reprint requests should be addressed.

Table 1 Properties of *E. coli* mutants defective in alkylation repair

Bacterial strains	Relevant genotype	Plating efficiency on nutrient agar containing 3 mM MMS* (%)	m ³ A DNA glycosylase activity in cell extracts (pmol base released per mg protein)			m ⁶ G DNA methyltransferase activity in extracts from induced cells (% m ⁶ G remaining in substrate)
			Noninduced (TagI)	Induced	Induced + 3 mM m ³ A (TagII)	
AB1157	Wild type	97	26.5	41.2	36.3	9.5
BK2106	<i>tag ada</i>	5×10^{-4}	6.4	0.7	1.8	94.5
BK2114	<i>tag</i>	67	3.8	26.3	34.0	6.5
BK2110	<i>ada</i>	1.4	24.3	23.1	5.3	94.0
MS23	<i>alk-1</i>	0.06	25.7	23.1	1.8	6.7
BD793	$\Delta alk(mgl-anP2H)$	0.16	27.0	26.2	0.5	2.9

Strain BK2106 is a MNNG-mutagenized derivative of AB1157 (ref. 9) and carries two mutations which both contribute to its MMS sensitivity. The mutation which maps at 47 min and co-transduces with *gyrA* is shown here to be in the *ada* gene and not in the *tag* gene as was previously assumed⁹. Other *ada* mutations have also recently been mapped to the same position on the map¹³. The *tag* mutation responsible for m³A DNA glycosylase (TagI) deficiency in BK2106 has now been mapped to a region between 70 and 74 min on the chromosome (G.E. and E.S., in preparation). Strains BK2106 and BK2114 (BK2106 *gyrA ada*⁺); BK2110 (AB1157 *gyrA ada*) and AB1157; and MS23 and AB1157 (refs 10, 11) are all isogenic pairs constructed by P1 transduction. BD793 was isolated by B. Duncan and is a *his*⁻ eductant of phage P2-lysogenized W1485F⁻. Cell extracts were prepared by a combination of plasmolysis and lysozyme digestion as previously described¹⁴. m³A DNA glycosylase activity was assayed by the method of Riazuddin and Lindahl⁸. The numbers presented are calculated on the basis of ethanol-soluble radioactivity released from ³H-dimethylsulphate-treated DNA using various amounts of extracts ranging from 0 to 20 µg protein. In this range a linear response was obtained between glycosylase activity and extract added. Parallel assays were also run where the reaction mixtures were subjected to polyethyleneimine TLC together with authentic 7-methylguanine and 3-methyladenine. In all cases where significant glycosylase activity was observed, both in induced and noninduced cells, >95% of the radioactivity released was identified as m³A. m⁶G DNA methyltransferase activity was assayed by the method of Karran *et al.*³ and 20 µg protein were added to each assay. Cells were induced by growth in M9 salts supplemented with 1% glucose and 1% casaminoacids in the presence of 0.5 µg MNNG for 90 min.

* Stationary phase cells calculated relative to plating on nutrient agar without MMS.

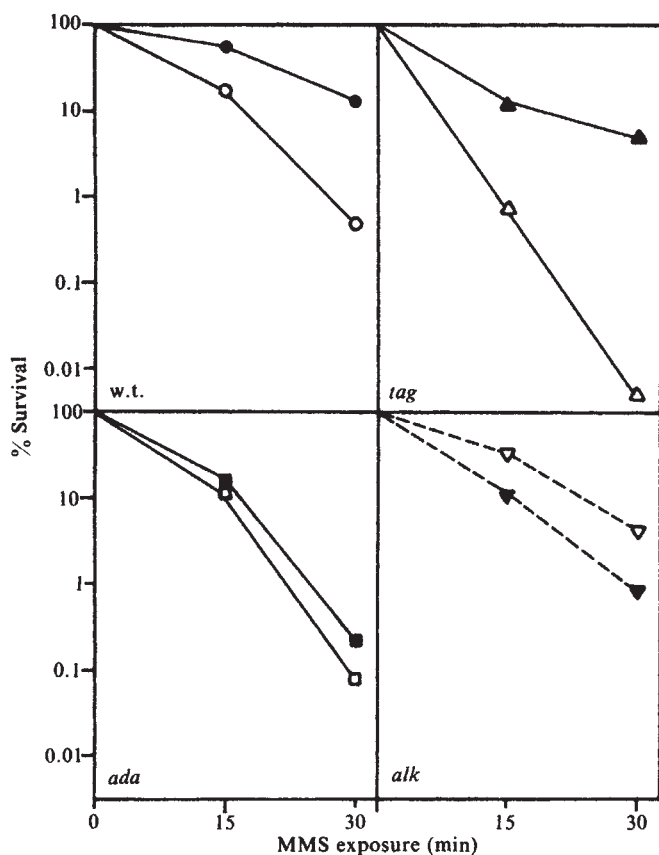


Fig. 2 Killing adaptation to alkylation resistance in wild-type and TagII-deficient mutants. Fresh overnight cultures were diluted 10-fold in K medium (M9 salts + 1% glucose and 1% casaminoacids). To one-half of each culture was added MNNG at a final concentration of 0.5 µg ml⁻¹ while the other half served as an untreated control. The cells were grown for 90 min at 37 °C and then washed once by centrifugation and resuspended in 1/10 volume of phosphate buffer. MMS was added at a final concentration of 0.06 M for strains AB1157 (wild type; ○, ●), BK2114 (*tag*; △, ▲) and BK2110 (*ada*; □, ■), and of 0.02 M for MS23 (*alk-1*; ▽, ▼). At intervals, samples were plated for survival after appropriate dilution. Closed symbols refer to adapted cell cultures (MNNG-treated) while open symbols represent untreated controls.

that TagII has a broader substrate specificity than TagI and is responsible for repair of a quantitatively minor but biologically important alkylation product other than m³A. TagI has a narrow substrate specificity and a very low *K_m* value for its substrate⁸, suggesting that it will quickly bind and remove m³A from alkylated DNA. Other minor products, however, could perhaps be removed by TagII. The *alk* mutants seem to have less residual TagII activity than the *ada* mutant, which would explain the greater sensitivity of the former compared with the latter (Table 1).

Mutants defective in the adaptive response were first described by Jeggo⁷. They were selected on the basis of their deficiency in mutagenic adaptation but most of them also appeared to be defective in killing adaptation. One mutant, however, was only defective in mutagenic adaptation. Jeggo *et al.*⁶ have further reported that *polA* mutants defective in DNA polymerase I are defective in killing adaptation, but show normal mutagenic adaptation. These studies are consistent with the view that mutagenic adaptation and killing adaptation are due to the induction of separate repair pathways. They support the view that m⁶G DNA methyltransferase is responsible for mutagenic adaptation since transferase repair of m⁶G will not require DNA polymerase I. The *ada*⁺-dependent inducible glycosylase may well be responsible for killing adaptation because DNA polymerase I is needed to complete repair of the apurinic site formed by the glycosylase action. A critical test for this interpretation is that *alk* as well as *ada* mutants should be defective in killing adaptation. Figure 2 shows that this is indeed the case and, furthermore, that the *alk-1* mutant is even sensitized by the adaptive treatment. These results therefore strongly support the view that TagII, not the transferase, is responsible for killing adaptation to alkylation resistance.

Thus *E. coli* seems to have two distinct m³A DNA glycosylase activities: one that is constitutively expressed and controlled by *tag*, and a second that is inducible and controlled by *ada* and *alk*. The constitutive enzyme appears responsible for rapid repair of m³A alkylation products in unadapted cells. The inducible enzyme seems to be required for killing adaptation to alkylation resistance and probably also for repair of some potentially lethal lesions not recognized by the constitutive enzyme in unadapted cells.

We thank Tomas Lindahl for communicating that strain BK2106 is defective in m⁶G DNA methyltransferase and for the gift of transferase substrate, Bruce Duncan and Mutsuo Sekiguchi for the gift of bacterial strains, Neil Clarke for valuable discussions, and Peter Strike for a critical reading of the

manuscript. This research was supported by the Norwegian Council for Scientific and Industrial Research.

Received 22 December 1981; accepted 24 February 1982.

- Samson, L. & Cairns, J. *Nature* **267**, 281–283 (1977).
- Cairns, J., Robins, P., Sedgwick, B. & Talmud, P. *Prog. Nucleic Acid Res. molec. Biol.* **26**, 237–244 (1981).
- Karran, P., Lindahl, T. & Griffin, B. *Nature* **280**, 76–77 (1979).
- Olsson, M. & Lindahl, T. *J. Biol. Chem.* **255**, 10569–10571 (1980).
- Schendel, P. & Robbins, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6017–6020 (1978).
- Jeggo, P., Defais, M., Samson, L. & Schendel, P. *Molec. gen. Genet.* **162**, 299–305 (1978).
- Jeggo, P. *J. Bact.* **139**, 783–791 (1979).
- Riazuddin, S. & Lindahl, T. *Biochemistry* **17**, 2110–2113 (1978).
- Karran, P., Lindahl, T., Øfsteng, I., Evensen, G. & Seeberg, E. *J. molec. Biol.* **140**, 101–127 (1980).
- Karran, P., Hjelmgren, T. & Lindahl, T. *Nature* **296**, 770–773 (1982).
- Yamamoto, Y., Katsuki, M., Sekiguchi, M. & Otsu, N. *J. Bact.* **135**, 144–152 (1978).
- Yamamoto, Y. & Sekiguchi, M. *Molec. gen. Genet.* **171**, 251–256 (1979).
- Sedgwick, B. *J. Bact.* (in the press).
- Seeberg, E., Nissen-Meyer, J. & Strike, P. *Nature* **263**, 524–526 (1976).

Binding of ATP to tubulin

James R. Zabrecky & R. David Cole

Department of Biochemistry, University of California, Berkeley, California 94720, USA

Microtubules are present in all eukaryotic cells and participate in a variety of cellular processes¹. From recent studies it has become clear that nucleotides have a central role in the assembly of tubulin into microtubules^{2–5}. The tubulin dimer binds two moles of guanine nucleotide: one (at the E-site) which can exchange with exogenous nucleotide and one (at the N-site) which does not exchange⁶. E-site GTP is hydrolysed during the assembly process^{4,7,8}. Microtubule assembly can be induced by ATP through a contaminating nucleoside diphosphokinase activity that regenerates GTP following its hydrolysis—this is associated with tubulin polymerization^{4,9,10}. Recent reports suggest that ATP has other effects on the assembly kinetics^{11–13}, and we have shown^{14,15}, using tubulin preparations lacking the nucleoside diphosphokinase activity, that ATP may play a critical part in the regulation of microtubule formation by binding to tubulin at a site which is distinct from the N- and E-sites. Kinetic data suggest a model in which ATP can act at the level of nucleation to stimulate assembly¹⁵. We now report the first direct evidence for the binding of ATP to tubulin and show that the corresponding dissociation constant is in a range of physiological significance.

In attempting to demonstrate the existence of a tubulin–ATP complex and to determine the dissociation constant, we found that equilibrium dialysis and centrifugation methods were unsuitable because they were too slow for such an unstable protein and because of the relatively weak binding of the nucleotide. We therefore chose the method of Hummel and Dreyer¹⁶ which is rapid and maintains the protein–ligand complex in equilibrium conditions. It has been used previously to determine guanine nucleotide binding to tubulin^{17,18}. In this method the protein is placed in solution with a given amount of ligand, and then applied to a gel filtration column which is equilibrated with the same concentration of ligand. As the protein emerges from the column at the excluded volume, it bears extra ligand above the equilibrium level, while at the included volume, the ligand level is depressed by the amount of ligand removed by the protein. The amount of ligand above the baseline level is equivalent to the protein–ligand complex and can be applied to the equilibrium expression, together with the concentration of free ligand, to calculate the dissociation constant, K_d .

Tubulin was prepared by chromatography on DEAE-cellulose (Whatman DE-52) as described elsewhere⁶. It has been shown to be more than 99% homogeneous electrophoretically and to have only minimal nucleoside diphosphokinase activity¹⁵. Figure 1 shows a representative profile obtained by passage of tubulin down a Sephadex G-25 column equilibrated with ³²P-ATP. As the protein eluted from the column, there was a corresponding increase in the level of nucleotide above

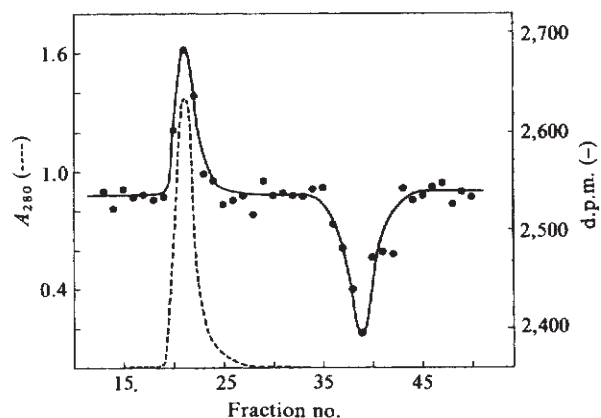


Fig. 1 Example of a Hummel–Dreyer column used to measure ligand binding. A Sephadex G-25 column (25 × 1 cm) was equilibrated with PM buffer (10 mM phosphate, 5 mM MgCl₂, pH 7.0) containing 20 μM [γ -³²P]ATP. Tubulin (~10 mg ml⁻¹) was prepared as described elsewhere^{5,10}, dialysed extensively against PM buffer, centrifuged at 140,000g for 30 min, then mixed with an equal volume of PM buffer with 40 μM [γ -³²P]ATP; 0.6 ml was applied to the column followed by elution with column buffer and 0.5-ml fractions were collected. Fractions were scanned for protein at 280 nm and the radioactivity determined by spotting aliquots on glass fibre filters (Whatman GF/C) and counting these using a Beckman LS7500 liquid scintillation counter. All chromatography was performed at room temperature.

the equilibrium value followed by a trough at the included volume. We ran five such columns for a series of ligand concentrations from 10 to 100 μM. Two peak fractions from each column were used to determine the level of ligand binding. The results are presented in Table 1.

The dissociation constant was calculated based on a stoichiometry of one binding site per tubulin dimer. The relative agreement in the values of K_d obtained at various ligand concentrations supports this assumption. The stoichiometry is usually determined from the amount of ligand bound in saturating conditions. However, with such a low binding affinity, the background levels of ligand would need to be so high in order to saturate the binding site that no reasonable precision could be obtained. Moreover, as relatively weak binding results in a rather modest difference between the peak of bound ATP and a substantial baseline, even at relatively low ATP concentrations, the average value calculated for the dissociation constant, 2.3×10^{-4} M, is only approximate. Clearly, however, ATP does

Table 1 Binding of ATP to tubulin measured by the Hummel–Dreyer method

Column ATP (μM)	Protein-bound ATP (μM)	Peak protein (μM)	mol ATP/mol protein	$K_d (\times 10^4)$
10	1.06	12.2	0.087	1.1
	0.665	7.4	0.090	1.0
20	1.17	12.8	0.091	2.0
	0.804	8.4	0.096	1.9
25	2.02	15.9	0.13	1.7
	1.20	10.2	0.12	1.8
50	2.10	10.0	0.21	1.9
	2.25	11.1	0.20	2.0
100	2.06	15.1	0.14	6.1
	1.87	8.5	0.22	3.5

Gel filtration columns equilibrated with the indicated concentration of [γ -³²P]ATP were run as described in Fig. 1 legend. The radioactivity in the two peak fractions was determined by liquid scintillation counting and the protein concentration by the method of Lowry *et al.*²⁰. K_d was determined from the equilibrium expression:

$$K_d = [\text{tubulin}][\text{ATP}] / [\text{tubulin-ATP}]$$

The stoichiometry was assumed to be one binding site per tubulin dimer (molecular weight 100,000).

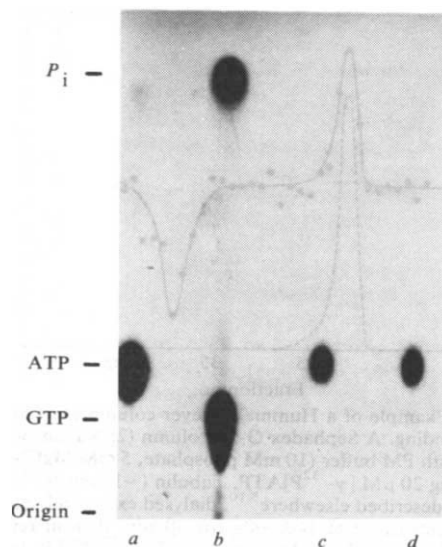


Fig. 2 Test for ATP hydrolysis following microtubule assembly. Tubulin ($\sim 40 \mu\text{M}$ in assembly buffer (10 mM 2-(*N*-morpholino)ethanesulphonic acid, 2.0 mM Mg^{2+} , 1.0 mM EGTA, 3.4 M glycerol, pH 6.4) with 1.0 mM GTP and 0.2 mM $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ was warmed to 37°C to allow microtubule polymerization. In these conditions, ATP significantly enhances the assembly kinetics¹⁵. Monitoring of the turbidity at 350 nm indicated that a steady state was reached within 10 min. Aliquots taken at 0 and 10 min were mixed with an equal volume of 95% ethanol, cooled to 0°C and centrifuged to remove the precipitated protein. These were spotted on to a PEI-cellulose thin-layer plate (Brinkman), then the plate was washed by immersion in 95% ethanol and dried. Ascending chromatography was performed using 1.0 M LiCl as the solvent, followed by autoradiography using Kodak X-Omat X-ray film. *a*, ATP standard; *b*, GTP and P_i standard; *c*, assembly sample (0 min); *d*, assembly sample (10 min).

bind to tubulin and the calculated K_d is within an order of magnitude of cellular ATP levels, that is, $\sim 2 \text{ mM}$ (ref. 19). It is therefore reasonable to suggest that ATP regulates microtubule formation *in vivo* according to changes in its distribution and concentration.

As ATP is bound to tubulin and affects the rate of microtubule formation it is appropriate to ask whether the ATP supplies energy through hydrolysis. Therefore, microtubules were assembled in the presence of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ and the solution was analysed by TLC for the presence of free phosphate. Although microtubule formation had reached a steady state, no free phosphate was observed (Fig. 2). (The appearance of a minor amount of GTP in the autoradiogram was due to a trace of residual nucleoside diphosphokinase activity in the tubulin preparation.) To exclude the possibility that the ^{32}P was incorporated into protein, the proteins in a similar microtubule mixture were resolved by electrophoresis and autoradiographed; no ^{32}P incorporation was detected. As ATP is hydrolysed neither to produce inorganic phosphate (P_i) nor to phosphorylate a protein during microtubule formation, the effect of ATP on assembly must involve some other mechanism, perhaps allostery.

This work was supported by research grants GMS 20338, EHS P30 ESO1896 and GMS 7232 from the NIH and the Agricultural Research Station.

Received 1 February; accepted 18 March 1982.

1. Roberts K. & Hyams, J. S. (eds) *Microtubules* (Academic, New York, 1979).
2. Weisenberg, R. C. *Science* **177**, 1104–1105 (1972).
3. Jacobs, M., Smith, H. & Taylor, E. W. *J. molec. Biol.* **89**, 455–468 (1974).
4. Penningroth, S. M. & Kirschner, M. W. *J. molec. Biol.* **115**, 643–673 (1977).
5. Weisenberg, R. C., Derry, W. J. & Dickinson, P. J. *Biochemistry* **15**, 4248–4254 (1976).
6. Weisenberg, R. C., Borisy, G. G. & Taylor, E. W. *Biochemistry* **7**, 4466–4479 (1968).
7. Kobayashi, T. *J. Biochem. Tokyo* **77**, 1193–1197 (1975).
8. David-Pfeuty, T., Erickson, H. P. & Pantaloni, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5372–5376 (1977).
9. Jacobs, M. & Caplow, M. *Biochem. biophys. Res. Commun.* **68** (1976); *Proc. natn. Acad. Sci. U.S.A.* **74**, 5372–5376 (1976).
10. Zeeberg, B. & Caplow, M. *J. biol. Chem.* **253**, 1984–1990 (1978).
11. Selkoe, D. J. *Brain Res.* **172**, 382–386 (1979).

12. Margolis, R. L. & Wilson, L. *Cell* **18**, 673–679 (1979).
13. Kumagai, H., Nishida, E. & Sakai, H. *J. Biochem. Tokyo* **85**, 495–502, 127–135 (1979).
14. Zabrecky, J. R. & Cole, R. D. *J. biol. Chem.* **255**, 11981–11985 (1980).
15. Zabrecky, J. R. & Cole, R. D. *J. biol. Chem.* (in the press).
16. Hummel, J. P. & Dreyer, W. J. *Biochim. biophys. Acta* **63**, 530–532 (1962).
17. Zeeberg, B. & Caplow, M. *Biochemistry* **18**, 3880–3886 (1979).
18. Caplow, M. & Zeeberg, B. *Archs Biochem. Biophys.* **203**, 404–411 (1980).
19. Williamson, D. H. & Brosnan, J. T. in *Methods of Enzymatic Analysis* (ed. Bergmeyer, H. U.) 2266–2302 (Academic, New York, 1974).
20. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).

Collective variable description of small-amplitude conformational fluctuations in a globular protein

Tosiyuki Noguti & Nobuhiro Gō

Department of Physics, Faculty of Science, Kyushu University 33, Fukuoka 812, Japan

Globular proteins in the native state undergo small-amplitude conformational fluctuations¹, which can be observed experimentally, for example, as temperature factors in X-ray crystallography^{2–5} or through their effects on the relaxation times in NMR⁶. The fluctuations can also be observed as the trajectories of each atom of a protein molecule generated by the theoretical method of molecular dynamics^{5,7,8}. These studies indicate that atoms or groups of atoms in the molecule, especially those near the surface, are in a 'fluid-like' state, which means that the forces acting on them have an impulsive character and the temperature factors obtained from X-ray crystallography have an anharmonic temperature dependence. More work is clearly needed to characterize better the state of the protein interior. Thus we have now calculated the shape of the conformational energy surface of a small protein, basic pancreatic trypsin inhibitor (BPTI), near its minimum. We find that for most degrees of freedom corresponding to collective conformational changes, the protein molecule behaves harmonically or like a solid.

In thermal equilibrium a protein molecule takes on a microscopic conformation with a probability determined by the con-

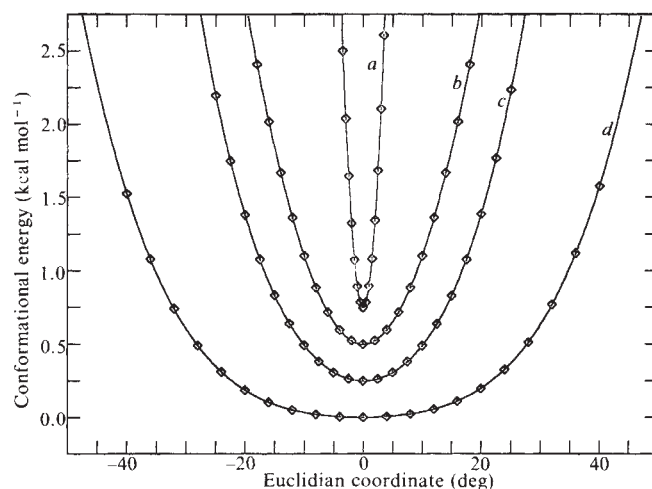
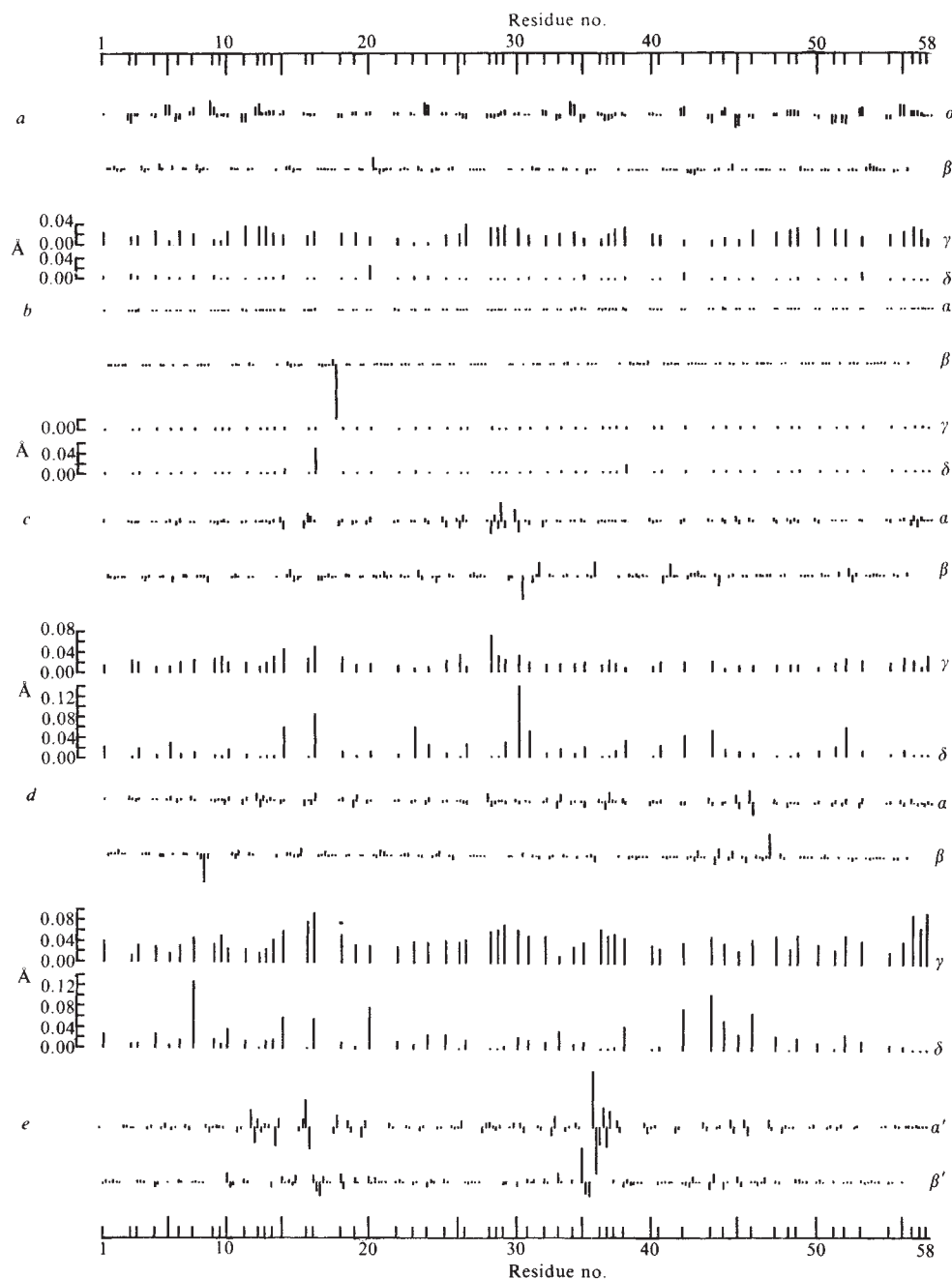


Fig. 1 Conformational energy curves calculated for actual conformational changes induced by variations in four variables, variables 37 (*a*), 118 (*b*), 179 (*c*) and 235 (*d*). The abscissa is the coordinate measured by euclidian distance defined as $d = \{\sum_{i=1}^{241} (\Delta\theta_i)^2\}^{1/2}$, where θ_i is a common notation denoting either ϕ or ψ or χ , and $\Delta\theta_i$ is the current value of θ_i minus that of θ_i at the energy minimum. Conformational energy values for *a*, *b* and *c* are off-set by 0.75, 0.50 and 0.25 kcal mol⁻¹ respectively for clarity. $\diamond$, Calculated conformational energy. —, Best-fit curves of the form of $B\xi^2 + C\xi^3 + D\xi^4$. Best-fit values of B (kcal mol⁻¹ deg⁻²), C (kcal mol⁻¹ deg⁻³) and D (kcal mol⁻¹ deg⁻⁴) for the four curves are: *a*, 0.15, 0.1×10^{-2} , 0.1×10^{-3} ; *b*, 0.0061, 0.3×10^{-7} , -0.6×10^{-6} ; *c*, 0.0023, 0.2×10^{-5} , 0.1×10^{-5} ; and *d*, 0.00032, 0.3×10^{-6} , 0.4×10^{-7} . The first three curves are essentially parabolas in the range of thermal fluctuations. Anharmonicity is observed in the last curve.

Fig. 2 Make-up of four eigenvectors of the second-derivative matrix, numbers 37 (a), 118 (b), 179 (c) and 235 (d), and equilibrium correlations of fluctuations in dihedral angles (e). In a–d, the value of the j th component, u_{ij} , of the i th eigenvector is represented by a bar (upwards when positive, downwards when negative) against serial number j of variable dihedral angles. When j refers to a main-chain angle, the bar is shown in lines α ; when to a side-chain angle, in lines β . Dihedral angles in the k th residue are arranged in the order $\phi_k, \chi_k, \chi_k^2, \dots, \psi_k$. Residue numbers, instead of the serial numbers of dihedral angles, are given on the abscissa. The displacements of atoms in collective conformational changes caused by changes in the collective variables are shown in lines γ and δ : γ , displacements of C'' atoms, and δ , average displacements of side-chain atoms relative to a local framework fixed to main-chain atoms in the same residue. The magnitude of the collective conformational change is taken as the root mean square thermal deviation of collective variable ξ_i calculated from the actual energy curve shown in Fig. 1. e, Example of relative equilibrium correlations of fluctuations in dihedral angles defined by $\langle \Delta\theta_i \Delta\psi_{35} \rangle / \{ \langle (\Delta\theta_i)^2 \rangle^{1/2} \langle (\Delta\psi_{35})^2 \rangle^{1/2} \}$. α' and β' refer to main-chain and side-chain dihedral angles. Residue 35 is located in an extended strand, which is roughly sandwiched between two other strands; it touches residue 11 in one strand and residue 17 in the other.



formational energy according to the Boltzmann law. The distribution of various fluctuating conformations can therefore be clarified by elucidating the shape of the conformational energy surface near the minimum point corresponding to the most stable conformation. Very near to the minimum, the energy surface is a multidimensional parabola characterized by the second-derivative matrix at the minimum.

The multidimensional parabola is most easily described in terms of a set of collective variables, $\{\xi_i\}$, corresponding to the eigenvectors of the second-derivative matrix

$$\xi_i = \sum_j u_{ij} \Delta\theta_j \quad (1)$$

where u_{ij} is the j th component of the i th eigenvector and $\Delta\theta_j$ is the deviation of the j th internal variable from that at the minimum point. A change in the variable ξ_i causes simultaneous and correlated displacements of many atoms in the molecule. Thus $\{\xi_i\}$ s are predicted to be natural variables describing conformational fluctuations in which strong correlations are known to exist among atomic displacements.

In terms of the collective variables, the conformational

energy, E , near the minimum point is given by

$$E = \frac{1}{2} \sum_i \lambda_i \xi_i^2 \quad (2)$$

where λ_i is the i th eigenvalue of the second-derivative matrix. For the range in which equation (2) applies, the collective conformational changes caused by changes in each ξ_i are mutually independent (that is, not correlated), and any conformational change can be predicted as superpositions of changes in ξ_i .

If the conformational energy can be expressed by equation (2), the conformational fluctuations will be harmonic or solid-like. Our previous study⁹ using a Monte Carlo method suggested that for many collective variables, the conformational energy was parabolic in the range of thermal fluctuations. Now we have calculated the second-derivative matrix at the minimum energy. Energy curves are calculated for a thermal excitation range of 0.3 kcal mol⁻¹ for actual conformational deformations caused by changes in each of the collective variables.

As in our earlier work⁹, bond lengths and bond angles in the protein are fixed and only dihedral angles of a main chain $\{\phi, \psi\}$ and side chains $\{\chi\}$ are treated as variables. The conformational

Table 1 Classification of anharmonic energy curves

Rank order of collective variables	Range of eigenvalues (kcal mol ⁻¹ deg ⁻²)	No. of anharmonic collective variables Type					Total
		I	II	III	IV	V	
1-53	1.30×10^{-4} - 1.00×10^{-1}						0
54-160	0.99×10^{-1} - 0.61×10^{-2}	3	1				4
161-180	0.60×10^{-2} - 0.45×10^{-2}				1		1
181-206	0.44×10^{-2} - 0.30×10^{-2}	2	6		1		9
207-229	0.29×10^{-2} - 0.10×10^{-2}	1	14	3			18
230-241	0.99×10^{-3} - 0.20×10^{-3}	1	3	5		2	11
Total		7	24	8	2	2	43

The energy curves for all the collective conformational changes are well fitted by fourth-order curves of the form $B\xi^2 + C\xi^3 + D\xi^4$, except for four curves classified as types IV and V. A curve is harmonic when the terms $C\xi^3$ and $D\xi^4$ are insignificant in the range of thermal fluctuations. As a measure of the range of thermal fluctuations, we take $[-\xi_0, \xi_0]$, where ξ_0 is defined by $\frac{1}{2}B\xi_0^2 = 0.3$ kcal mol⁻¹. A curve is anharmonic if $|C\xi_0^3/B\xi_0^2| \geq 0.1$ and/or $|D\xi_0^4/B\xi_0^2| \geq 0.1$, and may be one of five types according to its shape. Type I: curves where $|C\xi_0^3/B\xi_0^2| \geq 0.1$ and $|D\xi_0^4/B\xi_0^2| < 0.1$. In conformational changes caused by variations in these collective variables, large changes in dihedral angles and atom positions are confined to one or two side chains. II: curves where $1.0 > (D\xi_0^4/B\xi_0^2) \geq 0.1$. III: curves where $(D\xi_0^4/B\xi_0^2) \geq 1.0$. Conformational changes caused by variations in collective variables of types II and III are not localized in any part of the molecule. 4 of 24 type II curves and 5 of 8 entries in type III curves have non-negligible third-order terms, that is, $|C\xi_0^3/B\xi_0^2| \geq 0.1$. IV: conformational changes are localized in one or two side chains with smaller changes elsewhere. A peak of the energy curve is due to torsional energy for dihedral angle(s) of the side chain(s). V: Almost free rotation of a side chain (lysine) dihedral angle. Peaks are due to the non-bonded energy.

space therefore consists of 241 variables (116 main-chain and 125 side-chain dihedral angles). Hydrogen atoms are not treated individually but are included in calculations for the atom to which they are bonded, except for those expected to participate in hydrogen bonding. Geometrical parameters and energy functions are those used previously⁹, except for the energy functions for hydrogen bonding; energy functions of the form of (Ar⁻¹² + Br⁻¹⁰) are used. Solvation effects are not considered, that is, the molecule is treated as being *in vacuo*.

By starting from the conformation with the least energy in the previous Monte Carlo study, the conformational energy can be minimized in the 241-dimensional conformational space by the Newton method¹⁰ using first and second derivatives. The minimization is continued until a conformation is found in which all components of the first derivative are < 0.02 kcal mol⁻¹ deg⁻¹ and all eigenvalues of the second-derivative matrix are positive.

Figure 1 shows representative energy curves calculated for actual collective conformational changes. Serial numbers are given to the collective variables, corresponding to collective conformational changes in the descending order of the magnitude of the eigenvalues. For 198 of 241 collective conformational changes, the energy curves agree well with the parabola derived from the second-derivative matrix by the criterion described in Table 1 legend.

Deviations from the parabolas are observed in curves for the remaining 43 collective variables. All but five of them are members of the 61 softest variables (variables 181-241; see Table 1). We classified the 43 anharmonic collective variables into five types; 31 are such that the fourth-order terms have non-negligible positive contributions in the thermal fluctuations (see, for example, curve *d* in Fig. 1).

Thus, for many (~80%) of the collective conformational changes the conformational energy is harmonic. This strongly supports the use of collective variables to describe conformational fluctuations.

The nature of the conformational change induced by variation in the collective variable, ξ_i , is defined by the components, u_{ij} , of the eigenvector. The make-up of typical eigenvectors is shown in Fig. 2. Due to the complexity of the protein structure, there are many types of collective variable, most of which express non-local conformational changes (Fig. 2a, d). Whereas in variable 118, changes in dihedral angles and atom positions are localized mainly in one surface side chain (Fig. 2b), in variable 179, changes are localized mainly in a flexible loop structure consisting of residues 26-30 (Fig. 2c).

To characterize the shape of the multidimensional energy surface in the range of thermal fluctuations, contributions from higher-order cross-terms must be known. Even among the

harmonic variables these may be non-negligible. Some may cause anharmonic fluctuations such as transitions between energy minima. The anharmonicity observed in a few variables and higher-order cross-terms not considered here may account for the anharmonic behaviour of the conformational fluctuations observed in X-ray temperature factors²⁻⁵ and molecular dynamics^{5,7,8}.

There are two difficulties in our approach. First, the constraints on bond angles increased the energy barrier separating energy minima⁸. If the bond angles were to relax adiabatically, the conformational energy surface might be softer than we have allowed here, in which case the number of harmonic collective variables might be an over-estimate. Second, by neglecting the effects of solvent molecules, some changes in the surface of the molecule may not have been considered. However, our qualitative conclusion, that conformational fluctuations are largely harmonic and can therefore be described in terms of the collective variables, still holds. Because kinetic energy has not been considered, the collective variables defined here do not correspond to normal modes of vibration. The collective variables are useful in describing populations of conformations in equilibrium.

If we assume that, in the thermal range considered, the energy surface can be approximated in all degrees of freedom by a multidimensional parabola, we can calculate equilibrium correlations of fluctuations in dihedral angles from the second-derivative matrix (see Fig. 2e). Strong correlations exist between dihedral angles that are close in space, reflecting various collective conformational changes in the protein. If we can assume harmonicity of the energy surface, we will be able to calculate various quantities that reflect conformational fluctuations from the second-derivative matrix and to compare them with experimental results.

We thank Professor T. Ooi and Dr K. Nishikawa for providing a computer program to calculate conformational energy. Computation was done at the computer centres of Kyushu University and the Institute for Molecular Science. This work was supported by grants-in-aid from the Ministry of Education, Japan.

Received 9 November 1981; accepted 4 March 1982.

- Gö, N. *Proc. Japan Acad.* **56B**, 414-419 (1980).
- Frauenfelder, H., Petsko, G. A. & Tsernoglou, D. *Nature* **280**, 558-563 (1979).
- Artymiuk, P. J. *et al.* *Nature* **280**, 563-568 (1979).
- Sternberg, M. J. E., Grace, D. E. P. & Phillips, D. C. *J. molec. Biol.* **130**, 231-253 (1979).
- Northrup, S. H., Pear, M. R., McCammon, J. A., Karplus, M. & Takano, T. *Nature* **287**, 659-660 (1980).
- Ribeiro, A. A., King, R., Restivo, C. & Jandetzky, O. *J. Am. chem. Soc.* **102**, 4040-4051 (1980).
- McCammon, J. A., Gelin, B. R. & Karplus, M. *Nature* **267**, 585-590 (1977).
- McCammon, J. A. & Karplus, M. *Rev. phys. Chem.* **31**, 29-45 (1980).
- Noguti, T., Gö, N., Ooi, T. & Nishikawa, K. *Biochim. biophys. Acta* **671**, 93-98 (1981).
- Greenstadt, J. *Math. Comp.* **21**, 360-367 (1967).

MATTERS ARISING

Pleistocene and Holocene climates in Africa

IN examining the question of Pleistocene and Holocene climatic fluctuations in North-West Africa, Sarnthein *et al.*¹ distinguish between hypotheses that suggest distinct latitudinal shifts and those assuming stable latitudinal positions but contraction and expansion of the arid zone, and they cite Nicholson and Flohn² as an example of the former. This is clearly incorrect. We do indicate such shift at 18,000 yr BP, but not for later periods; and we clearly show a pronounced contraction of the arid zone during the climatic optimum at 6,000 yr BP. We do suggest that both a shift in the position and changes in intensity of the major circulation features are responsible for African rainfall changes in the late Pleistocene and Holocene.

Sarnthein *et al.*¹ cite primarily two lines of evidence of 'stable' latitudinal positions of circulation features. The first, concerning the location of the westerly wind belt over North-West Africa, is contradicted by recent findings³, which suggest it extended to 20–40° N at 18,000 yr BP and at present. Furthermore, the main Pleistocene displacement of that zone would probably have been during the summer and its southern most position (winter months) may not have differed much from present. The second piece of evidence is the mean position of the 'Harmattan' dust outbreaks, which are related to the mid-troposphere jet. The term 'Harmattan' is misused: today it is generally understood to be the surface, northeasterly trade-wind flow and not the upper-level easterlies. Furthermore, one cannot conclude, as those authors did, that the position of the ITCZ (intertropical convergence zone) is necessarily coincident with the position of the dust carried by the 'Harmattan' flow (or jet). This flow is a direct consequence of the temperature gradient between the Sahara and the relatively cooler and more humid air further south⁴. As long as the southern Sahara exists, and given the fixed Atlantic boundary to the south, the position of this easterly flow aloft (the mid-tropospheric jet) is relatively fixed. The ITCZ, on the other hand, is both dynamically and thermally determined and its position most directly reflects inter-hemispheric thermal differences, which did vary greatly throughout the late Pleistocene and Holocene. Although at present the ITCZ and related disturbances over West Africa coincide with the mid-tropospheric jet and Harmattan, they need not necessarily coincide.

The shear associated with the jet is critical for the formation of West African

disturbances⁴ and a weakened easterly flow could conceivably, as suggested by Sarnthein *et al.*¹, have been a factor in sub-Saharan aridity ~18,000 yr BP. Indeed, this may also have been a factor in the recent Sahel drought⁵. However, the authors present no evidence of a stable position of other circulation features, such as the westerlies or the ITCZ, and it is highly unlikely that the latter was not displaced towards the equator in late Pleistocene, when glaciation and temperature changes were most pronounced in the Northern Hemisphere and inter-hemispheric thermal contrast must have been minimal.

SHARON E. NICHOLSON

Graduate School of Geography,
Clark University,
950 Main Street, Worcester,
Massachusetts 01610, USA

1. Sarnthein, M., Tetzlaff, G., Koopman, B., Wolter, K. & Pflaumann, V. *Nature* **293**, 193–196 (1981).
2. Nicholson, S. E. & Flohn, H. *Climat. Change* **2**, 313–348 (1980).
3. Street-Perrott, F. A. & Roberts, N. in *Variations in the Global Water Budget* (eds Street-Perrott, F. A., Beran, M. & Ratcliffe, R. A. S.) (Reidel, Dordrecht, in the press).
4. Holton, J. R. *An Introduction to Dynamic Meteorology* (Academic, New York, 1979).
5. Kanamitsu, M. & Krishnamurti, T. N. *Mon. Weath. Rev.* **106**, 331–347 (1978).

TETZLAFF, SARNTHEIN AND WOLTER
REPLY—The comments of Nicholson give us the opportunity to clarify possible misunderstandings of our concepts and conclusions on the underlying causes of climatic change in North Africa¹. Nicholson particularly questions our evidence of 'stable' latitudinal positions of circulation features.

With regard to the extension of the westerlies, we basically agree with Nicholson, although we did not say too much about it in our article¹ and cited only indirect evidence. The westerlies indeed extended to 20–40° N over North-West Africa, in both periods considered, not only at 18,000 yr BP, but also at the present time. Nevertheless, average westerly wind vectors are substantiated on land only by proxy data² for precipitation due to synoptic disturbances in the westerlies.

In the Atlantic, the subtropical gyre and the coastal upwelling off the western Sahara essentially did not have a more southerly position around 18,000 BP and this provides additional evidence for a stable southern fringe of the westerlies during both winter and summer average seasons³. This contradicts a general equatorward displacement of the climatic belts south of the present westerlies in the Northern Hemisphere during cold stages, as suggested by Nicholson.

The other feature questioned—the relationship between the mean position of dust outbreaks and the position of the middle and lower tropospheric part of the ITCZ, in particular the monsoonal front—needs a deeper discussion of the meteorological structures. Along a trajectory in the continental trades southwards, the average temperature of the lower troposphere would continuously increase^{4,5}, at ~18,000 BP as well as at present. The middle tropospheric easterly winds which carry the main Saharan dust outbreaks require the opposite horizontal gradient. This means a southward sloping isobaric plane on the southern fringe of the subtropics. Like other authors⁶, we call these easterlies 'Harmattan' for our convenience.

We again agree with Nicholson that such an equatorward temperature decrease can be produced only by the cooler tropical air masses, the northern boundary of which is called monsoonal front and coincides with the ITCZ near the surface over land at the present. Finally, the maximum slope of the isobaric surface can also result in barotropic instability providing the energy for the easterly waves. This implies that the easterly (Harmattan) winds, the lower tropospheric parts of the ITCZ (Fig. 1a in ref. 1) and the easterly waves (including the dust transport in their wake) are three different effects of the same dynamic structure and their positions are necessarily closely related. We feel that some authors^{5,7} imply a connection between the positions of the maximum rainfall zone and the northern boundary of the monsoonal air masses. They are assumed to be coupled by the downdrafts of the storms in the rainfall zone. However, recent findings⁵ do not confirm such a direct relationship. Of course, the historical position of the ITCZ in North Africa is averaged over a number of cyclic processes on the synoptic, annual and long-term climatic scale, but it is not necessarily coupled with inter-hemispheric temperature differences as postulated by Nicholson. A separate dynamic ITC (intertropical convergence) equatorwards of the monsoonal front over Africa at 18,000 yr BP is theoretically possible, but not probable. As a result we consider our conclusions¹ valid. They strongly suggest that the dust-carrying winds, the Harmattan and the Trades, the subtropical anticyclone and also the ITCZ near the surface were not displaced equatorwards in western North Africa during the last glacial stage. Thus the fluctuations in the degree of sub-Saharan aridity/humidity should be related to factors such as the intensity and frequency of rain-bearing disturbances⁸ rather than to the average surface position of the ITCZ.

In our article¹, we did not intend to misinterpret Nicholson and Flohn⁹ when citing their hypothesis of distinct latitudinal shifts of the subtropical climatic belt from interglacial to glacial times. Accidental over-simplification of the alternatives may be one of the reasons for the apparent wrong citation. On the other hand, we feel rather corroborated in our statement by the comments of Nicholson discussed above.

G. TETZLAFF
K. WOLTER

*Institut für Meteorologie
und Klimatologie,
Universität D-3000
Hannover 21, FRG*

M. SARNTHEIN

*Geologisch-Paläontologisches Institut,
Universität D-2300 Kiel 1, FRG*

- Sarnthein, M., Tetzlaff, G., Koopmann, B., Wolter, K. & Pflaumann, U. *Nature* **293**, 193–196 (1981).
- Street-Perrott, F. A. & Roberts, N. in *Variations in the Global Water Budget* (eds Street-Perrott, F. A., Beran, M. & Ratcliffe, R. A. S.) (Reidel, Dordrecht, in the press).
- Sarnthein, M. et al. in *The Geology of the Northwest African Continental Margin* (eds von Rad et al.) 544–604 (Springer, Heidelberg, 1982).
- Dubief, J. *Le Climat du Sahara* Vols 1, 2 (1963).
- Wiin-Nielsen, A. & Krishnamurti, T. N. (eds) *Compendium of Meteorology* Vol. 2, Pt 4 (WMO no. 364, 1979).
- Jaenicke, R. & Schuetz, L. *J. geophys. Res.* **83**, 3585–3599 (1978).
- Pedgley, D. & Krishnamurti, T. N. *Mon. Weath. Rev.* **104**, 149–167 (1976).
- Houghton, D. in *Rep. U.S. GATE Central Program Workshop*, Boulder (1977).
- Nicholson, S. E. & Flohn, H. *Climat. Change* **2**, 313–348 (1980).
- Suzuki, K. & Mitsuoka, T. *Nature* **294**, 453–456 (1981).
- Waterhouse, J., Muir, C., Correa, P. & Powell, J. *Cancer Incidence in Five Continents* Vol. 3 (IARC Scient. Publ. no. 15, Lyon, 1976).
- Bonett, A. *Cancer in South Australia* (South Australian Government, 1980).
- Armstrong, B. & Doll, R. *Int. J. Cancer* **15**, 617–631 (1975).
- Drasar, B. S. & Irving, D. *Br. J. Cancer* **27**, 167–172 (1973).
- Phillips, R. *Cancer Res.* **35**, 3513–3522 (1975).
- Hirayama, T. *Nutr. Cancer* **1**, 67–81 (1979).
- Graham, S., Dayal, H., Swanson, M., Wittelman, A. & Wilkinson, G. *J. natn. Cancer Inst.* **61**, 709–714 (1978).
- Dales, L. G., Friedman, G. D., Ury, H. K., Grossman, S. & Williams, S. R. *Am. J. Epidemiol.* **109**, 132–144 (1979).
- Jain, M. et al. *Int. J. Cancer* **26**, 757–768 (1980).
- McMichael, A. J., Potter, J. D. & Hetzel, B. S. *Int. J. Epidemiol.* **8**, 295–303 (1979).
- McMichael, A. J. & Potter, J. D. *J. natn. Cancer Inst.* **65**, 1201–1207 (1980).
- Peto, R., Doll, R., Buckley, J. D. & Sporn, M. B. *Nature* **299**, 201–208 (1981).
- Baghurst, K. I. & Baghurst, P. A. in *The Assessment of Nutritional Status* (eds Hetzel, B. S. & Baghurst, K. I.) 139–160 (Menzies Foundation, Melbourne, 1981).
- U.S. Senate Select Committee of Nutrition and Human Needs, *Dietary Goals for America* (US Government Printing Office, Washington D.C., 1977).

Western diets and faecal nitrosamines

THE paper by Suzuki and Mitsuoka¹, while apparently demonstrating clear differences in nitrosamine production in individuals on different diets, makes unwarranted claims for the significance of these findings.

Japan has a low incidence rate of large bowel cancer compared with the United States² or Australia³. International correlation studies^{4,5}, studies of subcultures within societies⁶, and case-control and cohort studies of individuals^{7–10} have implicated dietary factors in colon carcinogenesis. Current hypotheses involve fat^{4,9,11}, fibre^{9,11}, alcohol¹¹, female sex hormones¹², vitamin A¹³, and cruciferous vegetables⁸ either as promoting colon cancer, via endogenous and bacterial bile acid metabolism, or as protective agents.

To test the hypothesis that a balanced Western diet constitutes a greater risk of exposure to nitrosamines (which act as initiating agents in some cancers) it is necessary to design an experiment which actually uses a balanced Western diet. The

diet used by Suzuki and Mitsuoka bears no resemblance to a normal Western diet. Specifically, the energy content of a typical Western diet^{14,15} is 9.7–10.5 MJ per day, compared with 11.8 MJ per day in the study of Suzuki and Mitsuoka. As a percentage of energy, protein intake is normally 12–15% (as opposed to 25%), fat ~40% (compared with 55%), carbohydrate ~45% (as opposed to 20%). In some European countries, certain vegetables may be consumed for breakfast but nowhere are salad vegetables eaten for breakfast. Indeed these vegetables were deliberately added to the 'typical' Western breakfast to increase nitrate intake. This ensures a good test of an hypothesis but not the one we are led to believe is being tested by the authors. Although Suzuki and Mitsuoka have established that a very high protein and fat diet with a large intake of nitrate increases the faecal output of nitrosamines, they have not demonstrated that this has anything to do with differences existing in the real world.

J. D. POTTER

*Division of Human Nutrition,
CSIRO, Kintore Avenue,
Adelaide, South Australia 5000*

- Suzuki, K. & Mitsuoka, T. *Nature* **294**, 453–456 (1981).
- Waterhouse, J., Muir, C., Correa, P. & Powell, J. *Cancer Incidence in Five Continents* Vol. 3 (IARC Scient. Publ. no. 15, Lyon, 1976).
- Bonett, A. *Cancer in South Australia* (South Australian Government, 1980).
- Armstrong, B. & Doll, R. *Int. J. Cancer* **15**, 617–631 (1975).
- Drasar, B. S. & Irving, D. *Br. J. Cancer* **27**, 167–172 (1973).
- Phillips, R. *Cancer Res.* **35**, 3513–3522 (1975).
- Hirayama, T. *Nutr. Cancer* **1**, 67–81 (1979).
- Graham, S., Dayal, H., Swanson, M., Wittelman, A. & Wilkinson, G. *J. natn. Cancer Inst.* **61**, 709–714 (1978).
- Dales, L. G., Friedman, G. D., Ury, H. K., Grossman, S. & Williams, S. R. *Am. J. Epidemiol.* **109**, 132–144 (1979).
- Jain, M. et al. *Int. J. Cancer* **26**, 757–768 (1980).
- McMichael, A. J., Potter, J. D. & Hetzel, B. S. *Int. J. Epidemiol.* **8**, 295–303 (1979).
- McMichael, A. J. & Potter, J. D. *J. natn. Cancer Inst.* **65**, 1201–1207 (1980).
- Peto, R., Doll, R., Buckley, J. D. & Sporn, M. B. *Nature* **299**, 201–208 (1981).
- Baghurst, K. I. & Baghurst, P. A. in *The Assessment of Nutritional Status* (eds Hetzel, B. S. & Baghurst, K. I.) 139–160 (Menzies Foundation, Melbourne, 1981).
- U.S. Senate Select Committee of Nutrition and Human Needs, *Dietary Goals for America* (US Government Printing Office, Washington D.C., 1977).

SUZUKI AND MITSUOKA REPLY—We agree with Potter that the diet used in our study does not exactly resemble a typical Western diet. To determine that nitrosamine formation in the intestine depends on diet, we used in our experiments a diet which contained a large

amount of nitrosamine precursors. Our study demonstrated that the faecal output of nitrosamines was increased by this kind of Western-style diet. However, we cannot implicate these diets directly in the aetiology of colon cancer. Further investigations of nitrosamine formation in the intestine in populations of high and low risk for colon cancer are necessary.

With respect to the p.p.b. analysis of nitrosamines in biological specimens, there are several problems associated with artefacts and contamination during the analysis. Nitrosamines are generally extracted from the specimens in alkaline conditions, because this inhibits nitrosation. We have noticed recently, however, that artificial formation of nitrosamines occurs in these conditions. This reaction seems to resemble the 'alkali effect' reported by Tozawa¹. We have used 1 mM of morpholine to monitor artificial nitrosamine formation during the analysis, as morpholine is nitrosated 500 times more readily than piperidine². However, we have recently found that morpholine is more difficult to nitrosate than piperidine in alkaline conditions, which renders morpholine unsuitable for use as a monitor of possible artificial formation of nitrosamines. We are re-investigating the faecal output of nitrosamines, thus our data³ may require correction.

KUNIO SUZUKI
TOMOTARI MITSUOKA

*Animal Physiology Laboratory,
The Institute of Physical
and Chemical Research,
Wako-shi, Saitama 351, Japan*

- Tozawa, H. *Bull. Japan Soc. scient. Fish.* **44**, 797–802 (1978); **45**, 1585 (1979); in *N-nitroso Compounds: Occurrence and Biological Effect* (eds Bartsch, H., O'Neill, I. K., Castegnaro, M. & Okada, M.) (IARC Publ. no. 41, Lyon, in the press).
- Mirvish, S. S. *Tox. appl. Pharma.* **31**, 325–351 (1975).
- Suzuki, K. & Mitsuoka, T. *Nature* **294**, 453–456 (1981).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

BOOK REVIEWS

Public rights, private land

Timothy O'Riordan

ANN and Malcolm MacEwen are an amiably dedicated couple who have spent the past three years on an exhaustive labour of love. They have visited each of the ten national parks of England and Wales, interviewed countless officers and members and consulted many other informed people, all with the aim of finding out whether the national park idea is alive and well in the land. To this important task they have also brought their own impressive knowledge of conservation and planning matters, plus Malcolm MacEwen's experience as an appointed member to the Exmoor National Park Committee. During that time he fought a partially successful campaign to protect Exmoor's magnificent moorland from being ploughed up and reseeded by landowners ready to capitalize from the grants and guaranteed prices available for intensive livestock production.

Their report is an important analysis of the current state of national parks in England and Wales. Together with another recently published document, *The Economy of Rural Communities in the National Parks of England and Wales* (1982), written by the Tourism and Recreation Research Unit at the University of Edinburgh, it provides by far the most comprehensive and sobering statement of the state of health of the national parks since the publication of the *Report of the National Park Policies Review Committee* (the Sandford Report; HMSO, 1974).

The MacEwens set out to discover whether the national park authorities are able to preserve for the nation the magnificent conservation and scenic heritage that caused the parks to be established in the first place. They believe that it is the duty of government (both central and at the national park level) to safeguard in trust for all citizens the many diverse scenic and natural history habitats which they claim have superior public rights to the private development rights of individual landowners. They are also very much aware that the national park landscapes are the product of care: they are not "natural" in the strict ecological sense, but a kind of semi-wild "garden", tended by generations of inhabitants to create a unique blend of cultural and

National Parks: Conservation or Cosmetics? By Ann and Malcolm MacEwen. Pp.314. Hbk ISBN 0-04-719003-5; pbk ISBN 0-04-719004-3. (George Allen and Unwin: 1982.) Hbk £15, \$35; pbk £8.50, \$18.95.

natural processes. The MacEwens call this characteristic "semi-wildness" and note that it has four vital elements: the opportunity for people to touch experience face to face; the repository both of natural and human history; the resource of space and openness; and a state where, until recently, man's relationship with nature has been relatively harmonious.

The MacEwens clearly regard these features as vital to the life-blood of the nation but realize that the future of the national parks cannot be separated from the future of the British countryside as a whole. Like the Edinburgh group, they are acutely aware that much of the national park land is harsh and economically inhospitable. The remote uplands are losing their working populations as agriculture is shedding labour, forestry is not attracting labour and the well-known problems of expensive and ill-provided transport lead to loss of job opportunities and services. Many of the national parks are suffering from the economic decline which is affecting remote rural Britain generally, and the local economies are so structurally weak that they cannot readily exploit the opportunities provided by tourism and the promotion of craft industries.

But the MacEwens' main criticism is levelled at the Ministry of Agriculture, Fisheries and Food (MAFF), the Common Agricultural Policy (and by implication, the Inland Revenue) for failing to recognize the social and conservation-related re-

percussions of policies favouring intensive agriculture and coniferous forestry. By means of a number of skilfully presented case studies they document how it pays landowners to exploit the public right of landscape in their own short-term private interest, aided and abetted by grants and tax incentives that positively encourage the destruction of the national park heritage. They further point out that the national park authorities are virtually powerless to stop this, lacking both the powers and the budgets to halt the decline. In addition, the MacEwens believe that the membership of national park committees should be less parochial and more committed to the original national park ideals.

But lest it be thought that the MacEwens are seeking a kind of landscape museum, one should make clear that they are anxious to extend public access to the semi-wild, and to develop the scope of environmental education and countryside interpretation so that many more people can enjoy the experiences that evidently mean so much to them. They are equally anxious to see a renaissance in rural economies favouring the retention of labour in agriculture and forestry so that key elements of the national park landscape can be retained. Indeed this is a critical point of their argument, and most relevant to any future government seeking to grapple with the problem of employment.

This means a major re-examination of MAFF policies and of the tax incentives favouring expenditures in social capital where this has both conservation and economic value. The MacEwens deplore the lost opportunity in the Wildlife and Countryside Bill debate when Lord Sandford's imaginative amendment, passed against the Government's wishes, fell in the Commons. This provision would

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

The price of farming? The broadleaved woodlands here in Patterdale, and elsewhere in the Lake District National Park, no longer regenerate naturally because of over-grazing.

have enabled MAFF to invest in rural enterprises to maintain employment in agriculture and related activities so as to promote conservation. The new clause is terribly weak, merely encouraging MAFF to further the cause of conservation but only within the narrow confines of their existing remit. The MacEwens are equally critical of the final shape of the Wildlife and Countryside Act with its inbuilt assumption that landowners are entitled to agricultural and forestry grants so that when such grants are denied on conservation grounds, the national park authorities will have to buy them out possibly to the tune of the profit foregone. They correctly emphasize that while these provisions are bad enough, the failure of the Act to cover the countryside outside the national parks and sites of special scientific interest could be disastrous. For here no notification is required of agricultural or forestry activity likely to be detrimental to conservation. It is hardly surprising that the agricultural community is anxious about its public image.

Like the Edinburgh group, the MacEwens conclude that the national park

authorities are toothless watchdogs, that their record has been more one of cosmetics than conservation and that the maze of bureaucratic interests with a finger in the countryside pie is emasculating coordinated action. They argue for a unification of the Nature Conservancy Council and the Countryside Commission (and possibly, eventually, the Development Commission?) on the grounds that the distinction between conservation and landscape is an unnecessary one, which may well lead to a continual weakening of the conservation cause.

There is little doubt that in the wake of the passage of the Wildlife and Countryside Act, the national park idea, already under fire because of changing economic circumstances in remote rural areas, is even more under threat. The MacEwens' arguments should be taken seriously or the nation is in danger of losing something that it will find very difficult to replace. □

Timothy O'Riordan is Professor of Environmental Sciences at the University of East Anglia. He is currently preparing the rural section of the UK response to the World Conservation Strategy.

The prosecution of syntax

John C. Marshall

To Mean — To Understand: Problems of Psychological Semantics. By H. Hörmann. Pp.337. ISBN 3-540-10448-8. (Springer-Verlag: 1981.) DM88, \$41.95.

MUCH of Professor Hörmann's latest book is taken up with attacks upon modern (that is, Chomskyan) linguistics and the psychological experimentation that has drawn upon Chomsky's ideas. Hörmann senses that the swings and roundabouts of outrageous fashion are deserting the paradigm that has dominated the field since 1957. "It is not uncommon", he writes, "that following a period of turbulent growth, a skeptical mood builds up and doubts begin to be expressed as to the rationale of progress achieved and the extent to which the position reached coincides with the original goals".

The particular malaise that Hörmann claims to detect has three main symptoms: current linguistic models are approaching "a level of complexity at which extreme sophistication borders on folly"; these models reveal a "widening gap between linguistic theory and the realities of language"; the last 25 years of work in linguistics and psycholinguistics has "contributed very little to a clarification of the problems of language use". Hörmann's antidote, as summarized in the last paragraph of the book, is that we should study how "the hearer is driven by the intentionality of his process of living-into-the-world", how "as he passes from one level to the next, the sounds, words, and

sentences of the language become transparent" and "fade away to make room in his consciousness for the meaning meant".

I have no doubt that, sociologically, Hörmann's remarks are exactly to the point and that both his diagnosis and his remedy will strike a responsive chord in the breasts of many psychologists. Intellectually, however, the argument is more dubious. Consider first the charge of over-complexity. No one would deny that, *ceteris paribus*, simple grammars should be preferred to intricate ones, but it is nature, not the linguist, that determines how elaborate the grammars of natural language are. And in any case, Hörmann seems unaware that a radical simplification of linguistic theory has taken place over the last decade; the proliferation of rules and rule-types that characterized generative syntax in the mid-1960s has been replaced by a system of very general principles whose interaction is responsible for the appearance of "phenotypic" diversity.

In order to appreciate the second complaint one must consider what "the realities of language" entail. For Hörmann,

human language is language because it is used by people for a purpose, namely, to live with other people. The purposeful use of language is embodied in acts of meaning and of understanding; in these acts the essence of language is integrated with the condition of man.

While the passage may be a trifle purple, the sentiments are surely unimpeachable. But why does Hörmann believe that gener-

ative linguists have ignored "meaning" and "understanding"? The central themes of Chomsky's work have always included the grammatical representation of thematic structure (who did what to whom with what), the referential possibilities of anaphoric elements, and the interpretation of the logical vocabulary of natural languages. There may be more to meaning than this, but surely these topics are part of "semantics" in anyone's story?

Hörmann's final point, that the study of language-use has not been advanced by recent psycholinguistic research, seems at first blush a particularly odd way of describing a period that has seen notable advances in modelling sentence parsing and production. Apparently, that area is not what Hörmann refers to as "language-use". His fulminations are rather based upon the failure of generative linguistics and psycholinguistics "to predict events as they occur in everyday life". Hörmann's claim may be confidently conceded, although the complaint smacks of expecting Newton's laws of motion to predict when the last apple will fall from the tree in my back garden.

The bone of contention is this: common sense and Professor Hörmann know that there's a lot going on when two people talk with each other. In addition to speaking a common language (English, German, Malagassy, or whatever), the participants also bring to bear their tacit grasp of conversational maxims, and their vast knowledge of "the way the world is". Professor Hörmann notes, correctly, that theories of generative grammar do not explain everything about these "acts of meaning and understanding". But he then concludes that such theories therefore explain nothing. This hasty dismissal rather misses the point that however much extra-linguistic factors may aid in understanding what someone *really* means, the greatest help surely comes from knowing the other person's language. It is precisely this knowledge that grammars explicate. The scope of formal grammars may, however, be very restricted. For example, Chomsky's position is that a particular level of linguistic representation — "the syntax of logical form" — is the interface between sentence-grammar and other faculties of mind implicated in the wider sense of the weasel word "semantics". Beyond the level of logical form no sharp dividing line is drawn between the "literal" meaning of a sentence and its interpretation in the light of the beliefs, knowledge and pragmatic inferences that can be brought to bear upon it. It is, of course, an empirical issue whether mental representations do indeed partition in this fashion. The pity is that Professor Hörmann obscures these topics by arguing that sentence-grammars fail to solve problems they were never intended to solve. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford.

Littermates and guesstimates

Barry Cox & Michael Stoddart

The Mammalian Radiations: An Analysis of Trends in Evolution, Adaptation, and Behaviour. By John F. Eisenberg. Pp.610. US ISBN 0-226-19537-6; UK ISBN 0-485-30008-7. (Chicago University Press/Athlone: 1981.) \$45, £32.

THE theme of John Eisenberg's book is the relationship between three aspects of mammals: their anatomy and physiology, the way in which they exploit their different habitats, and their behaviour and social structure. Clearly, these aspects must be integrated in any species, but Eisenberg goes further in attempting to compare the patterns of integration in different mammalian groups, and to suggest general trends that extend from one group to another.

The first 74 pages deal with the relationship between the areas of origin of the different groups and the changing continental configurations, and survey what Eisenberg calls "the early radiations" — monotremes, marsupials, edentates and lemurs. In the next 150 pages he deals with the remaining orders of mammal in similar fashion. He is at his least authoritative in the historical and taxonomic section, for he shows the Marsupialia as already in existence in the Lower Cretaceous, while the suggestion that monotremes evolved from multituberculates that reached Australia is to found one highly dubious hypothesis upon another. After uniting the South American ungulates with such African forms as elephants, Eisenberg notes that this suggests that they had a common ancestor in those areas before the two continents separated — though the separation took place in the mid-Cretaceous, long before the differentiation of the ungulate eutherians. Furthermore, are the marsupials and edentates simply part of an "early radiation" of mammals? Since Eisenberg is attempting to deduce phyletic trees of behavioural traits, and these trees are in turn founded upon the classification and history derived from morphological studies of fossil and living forms, it is essential that he is accurate in these fundamental assumptions.

Eisenberg's method and assumptions are clearly explained on p.12.

From the initial radiations, there are today some living mammals, mostly in the tropics, that in many morphological features show relatively little modification from the original stocks . . . it seems legitimate to infer that certain morphologically conservative species inhabiting niches in the tropics represent forms that are in a sense adapted to a 'conservative niche'. Thus a phylogenetic reconstruction of mammalian behavioural evolution should start with generalizations developed by comparing morphologically conservative forms that occupy what might be considered the niches of the Paleocene.

He goes on to comment (p.42) that "The

age of the edentate and pholidote lineages allows them to serve as a basis for comparison with the remaining radiations of the eutherians". Finally (p.49)

If the Edentata are among the most primitive living eutherians, then the behaviour of a 'conservative' dasypodid such as *Euphractus* has great significance for comparisons with other morphologically conservative mammals. Unfortunately we have virtually no field data . . . We must therefore rely heavily on captive studies.

The gaps in this chain of reasoning are obvious. To what extent is any niche today equivalent to a niche in the Palaeocene? Are armadillos really morphologically conservative? If they are, at least in those characters that condition their ecological niche, to what extent does this allow us to deduce that they are also conservative in their reproductive and behavioural strategies? How much meaning is there in the concept of a "phylogenetic reconstruction of mammalian behavioural evolution"? These are important reservations, that must be raised now because they counsel caution in reading the rest of the book.

In Chapters 17–23, Eisenberg discusses what he terms "macrophysiology and adaptation" — an integration of dietary specializations and ways of life. This is the classical stuff of mammalian adaptive radiation, and his approach is useful and thought-provoking. For example, why are fossorial niches incompletely occupied in the tropics, and why are more "macroniches" filled in North America than in Europe or North Asia?

In the final 100 pages of text, Eisenberg develops the theme that there was originally a common basic mammalian behavioural repertoire, which has been adapted through natural selection in the

various diversifying lineages. He first attempts to identify this repertoire by comparing various patterns of behaviour in terrestrial mammals; this chapter is brief and rather unsatisfactory. However, in a detailed and most valuable analysis of social organization, Eisenberg spiritedly tries to correlate such behavioural systems as mating, rearing, foraging and refuting. He then extends his net to show relationships between demographic, ecological and sociological variables in a wide range of species. The use of arbitrary numerical values for qualitative characteristics (e.g. "reproductive rate" is scored from 1 to 7) is understandable but can be misleading. Many species vary their reproductive output and adopt a slightly different social organization at different altitudes or latitudes, so the reader must beware of these oversimplifications. However, the author seems more at home in this section than in the earlier parts of the book, and his extrapolations should fuel many a coffee-room debate.

There is an excellent and wide-ranging bibliography of over 1,700 references, and Eisenberg has had the happy idea of entering many of these in group versus subject tables that show, for example, references on rhinoceros behaviour or tapir ecology. Appendices, classifications of extant mammals and an index complete the book.

For research workers and postgraduate students this is certainly a stimulating book, full of data, correlations, extrapolations and provocative guesses. It is therefore dangerous material for the unwary or inexperienced, who may be unaware of the fundamental assumptions that Eisenberg's thesis is founded upon, and may be slow to appreciate the boundaries between correlation, extrapolation and guesswork. □

Barry Cox is a Professor and Michael Stoddart a Lecturer in the Department of Zoology, King's College, University of London.

Life of the long-haired stars

David W. Hughes

Introduction to Comets. By John C. Brandt and Robert D. Chapman. Pp.246. ISBN 0-521-23906-0. (Cambridge University Press: 1982.) £21, \$45.

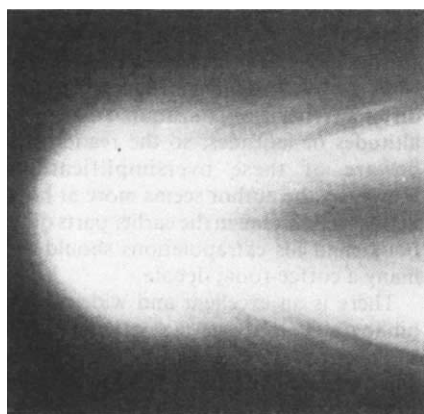
COMETS are heavenly bodies of considerable beauty, volatile character and deep physical and chemical significance. Their contents are fresh, thermally unaltered and undifferentiated, and the large outpourings of gas and dust that occur when the comet is close to the Sun often produce a coma and tail of truly astronomical proportion. Comets are also newsworthy. The past decade has seen the

blossoming of ultraviolet and infrared comet research; the next will see the launching of the first spacecraft designed to pass through the head of a comet and to gather *in situ* data.

Monographs on comets are scarce and there have been only three major ones since the turn of the century: G.F. Chambers's *The Story of the Comets* (Oxford University Press, 1910); R.A. Lyttleton's *The Comets and their Origin* (Cambridge University Press, 1953); and N.B. Richter's *The Nature of Comets* (Methuen, 1964 — an enlarged translation of his 1954 book, *Statistik und Physik der Kometen*). We

now have Brandt and Chapman's *Introduction to Comets*. Let me say straight away that a very good introduction it is too; a worthy and timely fourth addition to the list. The authors have conveyed the excitement of modern comet work, the pace of the data flow and the originality of many of the new theoretical approaches. They are excellent recruiting agents for the subject.

The book starts by placing comets in the historical perspective and reviews many of our modern ideas about comets by discussing how scientists first approached, for



Halley's head, photographed on May 8, 1910, from the Hale Observatory.

example, cometary spectroscopy and the interactions of comets with the solar wind. The second section deals in depth with the dynamics and structure of comets and looks at the models that have been proposed for the chemical reactions and dust emission processes. The origins of comets and their decay into meteor streams are also considered. The third section reviews the enormous advances that have come from the investigation of some of the recent bright comets, such as Kohoutek, Kobayashi-Berger-Milon and West. It also discusses the possible spacecraft missions to comets. The last section briefly touches on comet germs and impacts, and the place of comets in literature and art.

The book is extremely well illustrated and delves into the nature of comets with sufficient detail to satisfy the average undergraduate and postgraduate reader. Those seeking more, are efficiently led to an impressive list of publications in the "suggested reading" section.

I think the authors have used their 246 pages most effectively. For my taste I would have put in more than three pages on origin mechanisms, and would have laid more emphasis on orbital evolution, decay process and size variation. Also the section on cometary missions seems to have been written long before the publication date, in the rosy times when NASA were still actively planning cometary missions. The people who *are* going to Halley in 1986 don't get a mention. □

David W. Hughes is a Lecturer in Physics and Astronomy at the University of Sheffield.

Progress and prospects in solar physics

J. C. Brown

Solar Active Regions: A Monograph from Skylab Solar Workshop III. Edited by Frank Q. Orrall. Pp.350. ISBN 0-87081-085-5. (Colorado Associated University Press: 1981.) \$17.50.

IDEALLY the trilogy of Skylab Solar Workshop proceedings would have progressed from the tranquillity of *Coronal Holes* through *Active Regions* to achieve its nexus in *Solar Flares*, thereby exhibiting the physical progression of solar phenomena. In reality even the comparative opulence of workshop funds, albeit a tiny fraction of mission costs, did not permit this degree of coordination. Nevertheless the substantial overlap of participant lists means that the three workshop volumes together contain useful interdisciplinary seeds for the solar activity student to cultivate. Benefiting perhaps from previous experience, this third volume is the most readable.

The editor opens with a thorough overview of workshop results from each of the five teams who contribute about two detailed chapters each. A recurring theme throughout is theoretical and observational modelling of static loops of which the corona is now held mainly to consist. It appears to be generally accepted that spectral analysis can shed only dim light on details of loop heating mechanisms, though scaling laws permit comparison of the magnitude of heating in relation to loops of different lengths and pressures. Aside from these static analyses, the work includes a lengthy theoretical discussion of gas flows and thermal instability in loops, complemented by two chapters on observations of dynamic phenomena at chromospheric/coronal levels and near the photosphere, the latter invoking data from OSO-8 as well as from Skylab.

For the non-specialist the chapters of greatest interest will be those on theoretical coronal heating mechanisms and on spectral diagnostics, both of which are lucidly presented. The former contains much on energy release in plasmas which will interest plasma astrophysicists in general, while the latter concisely reviews the principles of spectrum line diagnostics for plasma temperatures, densities and differential emission measures. A frank assessment of progress made and of prospects for future work on active regions completes the monograph.

As usual this workshop raised more questions than it resolved, reiterating the stark solar message for cosmic astronomers of just how difficult astrophysical problems become when the object of study is close enough for detailed scrutiny. The Skylab Workshop series presented a unique opportunity for thrashing out both theoretical and observational progress in

important specialized problems. It is to be hoped that the sorry current state of NASA funding does not preclude such in-depth attacks on data from more recent missions such as Solar Maximum. □

J. C. Brown is Reader in Astronomy at the University of Glasgow.

Sex differentiated

W.K. Whitten

Mechanisms of Sex Differentiation in Animals and Man. Edited by C.R. Austin and R.G. Edwards. Pp.603. ISBN 0-12-068540-X. (Academic: 1981.) £37, \$89.

IN 1970, Harris and Edwards edited the Royal Society's discussion *Sex Determination*, and in 1972 a chapter by Short entitled *Sex Determination and Differentiation* appeared in *Reproduction in Mammals*, edited by Austin and Short, and published by Cambridge University Press. Now Austin and Edwards restrict their title to *Sex Differentiation* and argue that determination is the initial step in homotypic differentiation. But this is not accepted by some of their contributors, so the editors conclude that the argument is unresolved — a healthy attitude, emphasizing the need to standardize terminology.

This volume of 12 chapters written by specialists is intended as a text for advanced students and researchers. Lyon, Cattanaach and Charlton give superb descriptions of testicular feminization, XY female wood lemmings, and sex-reversed and hypogonadal mice. The study of these conditions and the H-Y antigen, described by Wachtel and Koo, have contributed greatly to the recent explosion in our understanding of sex differentiation. There are excellent chapters on embryonic development and function, postnatal growth of gonads and sex characters from the clinical point of view, and the many defects in determination and differentiation in human beings. The structure of the gonad, however, is followed only as far as initiation of meiosis. I assume that this is because folliculogenesis and gametogenesis are covered in other recent books such as Peters and McNatty's *The Ovary* published by the University of California Press in 1980. There is a review of X-chromosome inactivation, but unfortunately no contribution by Ohno who anticipated many of the recent advances.

Benirschke adequately reviews freemartins and hermaphrodites of large animals, but his treatment of mosaics and chimaeras of mice is poor. Apparently he

has not read Russell's Oak Ridge Symposium volume, *Genetic Mosaics and Chimeras in Mammals* (Plenum, 1980), nor the more recent work with fetal mouse hermaphrodites. These topics really deserved a chapter by Beamer whose BALB/CBea mice generate mosaic hermaphrodites by nondisjunction, or Eicher who has transferred the Y-chromosome of *Mus poschiavinus* to C57BL/6J-Y^{POS}. This chromosome carries a variant of the testis-determination locus which interacts abnormally with autosomal or X-linked loci preventing normal testis differentiation even in the presence of H-Y antigen. Also there are excellent controls from consomic strains for these interesting animals.

The words "mechanisms" and "animals" in the title are misleading because some chapters, such as Stevens's account of his exciting spontaneous and induced teratomas, are purely descriptive, and except for a few pages of introduction and an excellent contribution on environmental and non-genetic mechanisms the book is restricted to mammals.

Reviewing this book has been difficult because many of the references cited, including books, are recent and not freely available in libraries with restricted budgets. There is no author index and the inadequate subject index appears to have been compiled from the table of contents. An editorial policy of giving the specific name when a species is first mentioned would have ensured entries for "man" and "mouse", at least in the species index, and eliminated some redundant ones. Surely textbooks warrant the services of professional indexers?

The editors consider the possibility of voluntary control of sex determination and doubt the "dire prognostications" that acquisition of this knowledge would upset the "fabric of society". Instead, they predict smaller, planned families and elimination of sex-linked disease, but conclude that no foolproof method exists at present. Already blastocysts can be sexed before implantation and fetuses of the unwanted sex aborted. Ironically, the early data from test-tube babies gave an excess of females ($P < 0.03$) which might have been interpreted to mean that a procedure favouring girls had been stumbled upon. But I believe that biologists generally should require a probability of less than 0.01 before assigning significance, and this conviction has been supported by more recent results, particularly from Melbourne.

In spite of these criticisms I consider this a valuable and major work, updating Crew's little gem *Sex Determination*, published by Methuen in 1965, and fit to stand alongside White's *Animal Cytology and Evolution* (Rand McNally, 1973). □

W.K. Whitten is a Research Associate in the Department of Zoology, University of Tasmania.

News of conception, pure and applied

C.R. Austin

Fertilization and Embryonic Development in Vitro. Edited by Luigi Mastroianni and John D. Biggers. Pp.371. ISBN 0-306-40783-3. (Plenum: 1981.) \$45, £28.35. *In Vitro Fertilization and Embryo Transfer*. Edited by E.S.E. Hafez and K. Semm. Pp.393. UK ISBN 0-85200-438-9; US ISBN 0-8451-3005-6. (MTP Press/Alan R. Liss: 1982.) £29.95, \$58.

THE recent upsurge of attention paid in the public media to the growing number of normal childbirths following the fertilization of human eggs *in vitro*, and the culture and reinsertion of the resulting embryos into their "mothers", must serve as welcome publicity for these two books, both of which provide in their own way the background to those newsworthy events. The publicity is a late development, for the vigour of scientific research in this field has been building up for the past 25–30 years, and the intellectual springs behind this drive are based on biologists' ambitions that go back a century or more.

Some of the history is discussed in the Prologue to the book edited by Mastroianni and Biggers, though this account is concerned primarily with milestones along the trail of the recovery and culture of mammalian embryos. From early times, investigators have been fascinated by the prospect that conditions might be achieved in the laboratory permitting the union of mammalian gametes and the ensuing development of the embryos in such a way that these events could be held under direct observation *in vitro*. Also early in the story, the remarkable resilience of mammalian embryos to experimental manipulation attracted devoted study, and the first successes with the transfer of embryos from one animal to another were achieved before the turn of the century. With increasing sophistication the focus has sharpened in the modern era so that investigators have become concerned more and more with fundamental details.

The two books under review have, not surprisingly, much in common. They are of about the same vintage, with that edited by Hafez and Semm marginally the older as it is based on papers presented at a world conference in 1980. The other is a "planned" book of 15 chapters (21 authors, all in American laboratories), while Hafez and Semm is comprised of 37 papers (71 authors, of whom 51 work in countries other than the USA, chiefly West Germany, France and Japan). The common ground of subject matter includes the recovery and preparation of gametes, *in vitro* fertilization, the culture and evaluation of embryos, and embryo transfer; differences arise from the manner in which the subjects are handled and in the identity of peripheral topics. Broadly

speaking, Mastroianni and Biggers is biased towards basic problems and Hafez and Semm towards the applied.

The former volume is distinguished by a monumental chapter by R. Yanagimachi on mechanisms of fertilization in mammals; this, together with three others (by G. Oliphant and L. A. Eng, B. D. Bavister, and D. P. Wolf), tells the reader almost all there is need to know about sperm maturation and capacitation, the acrosome reaction, penetration of egg investments, sperm-egg fusion, pronuclear life and the blocks to polyspermy — all this in laboratory animals. *In vitro* techniques are stressed. In addition, this book offers two chapters on gamete interactions and egg activation in the sea urchin (a worthy comparative background, though inessential in the present context), two chapters on chromosomal anomalies affecting embryonic development, two more on the rather esoteric topic of water and electrolyte balances in blastocysts, and finally a well planned and informative chapter on embryo transfer in cattle.

In Hafez and Semm, by contrast, no less than 14 papers describe the application of methods for *in vitro* fertilization, and four deal with embryo transfer. About half the papers in this book accord at least a mention to the human subject, for whom the timing of ovulation, oocyte recovery and maturation *in vitro*, fertilization *in vitro* and embryo evaluation are treated in varied but generally satisfactory detail, and the ethics of embryo transfer are briefly discussed. Curiously, none of the research groups that have reported full success with human material — well established pregnancies and normal births — are represented here. Other papers are concerned with non-human primates and farm animals. The book also fortunately provides us with Pierre Soupart's report, bringing us up to date with his work on initiating development by the singular device of fusing two oocytes; this may have been his last paper. In addition, there is a report on gonadotrophin production by baboon embryos *in vitro* from V.Z. Pope *et al.*, and descriptions of two recently developed rapid assay methods for LH by M. Taymor *et al.* and by R. Frydman *et al.*

Both of these books are mines of information and "musts" for all gametologists; if only one is to be purchased the choice must depend on the predilection of the reader — whether for the basic or the applied, for the transatlantic or the international, or simply for the more or less expensive. □

C.R. Austin is Emeritus Charles Darwin Professor of Animal Embryology, University of Cambridge.

In search of the sublime

S. M. Walters

The Garden of Eden: The Botanic Garden and the Re-creation of Paradise. By John Prest. Pp.122. ISBN 0-300-02726-5. (Yale University Press:1982.) \$25, £12.50.

AMONGST the ever-growing literature on the history of gardens, relatively little explicitly treats of botanic gardens and their place in the general development of horticultural ideas and practices. It is therefore of some interest to note the subtitle of John Prest's book, and to see how far the text which accompanies the lavish illustrations lives up to the promise. The thesis of the book is interesting, even novel: he traces the idea (or more strictly *ideas*) of Paradise in Christian thought and imagery as reflected in the design of gardens through the centuries, and places the sixteenth and seventeenth century botanic gardens firmly in this tradition. Whilst it is easy to agree that Christian images and influences were powerful in gardening history, the thesis does not wholly convince, for it underestimates the scientific and instructional aspects of the work of botanic gardens.

As befits an Oxford historian, the author has chosen the oldest botanic garden in England, that in his university, to illustrate the shape and scope of early botanic gardens, but it is surely too parochial altogether to ignore Kew except for a single, oblique reference. The final chapter, which discusses the breakdown of isolation of the *hortus conclusus* in the new climate of interest in nature and the natural world, is surely the place where Kew Gardens cries out for treatment. The interplay of botanic gardens with the great private gardens of the nobility is indeed a fascinating theme, but it is curiously and disappointingly undeveloped in this book.

The relationship between the plants and the animals in the Garden of Eden — one of the splendid illustrations in colour (also used on the dust-jacket) depicts this scene painted by Jan Brueghel — is touched on throughout the book, with the strong implication that it was only the sheer impracticability of including animals that prevented botanic gardens from being also zoos. This is surely a serious distortion of the history of botanic gardens, which arose largely to satisfy the needs of medical teaching in the universities. Although it is true that botanic gardens later took on the general function of displaying as wide a range of plants as possible, there is no evidence that they were ever seriously troubled by the vision that they might also display a correspondingly wide range of

representatives of the animal kingdom. Why should they, when systematic zoology was not taught in the universities, either to medical students or to anyone else?

It is perhaps inevitable that a book written by a historian on botanic gardens should say very little about the plants themselves, although they are the *raison d'être* of the institutions which grow them; when we find, however, that such few Latin scientific names as are used are mostly given incorrectly, with no capital letter for the generic name (as on pp.61 and 83), our confidence in the author's knowledge of this part of his subject is seriously undermined.

Undoubtedly, the author has read both widely and deeply in and around his subject, as the many footnotes, sensibly collected at the end together with the excellent bibliography, bear witness. As a source for further exploration of interesting themes, the book is quite exceptionally useful. The quality of paper, printing and reproduction of illustrations is very high, and the price, by 1982 standards, quite modest. A minor irritation is that many of the pages are unnumbered, because of the placing of illustration captions: this could surely have been avoided, without detriment to the very attractive design. □

S. M. Walters is Director of the University Botanic Garden, Cambridge, and author of *The Shaping of Cambridge Botany* (Cambridge University Press, 1981).

Chemical soul

A.G. Lee

Liposomes: From Physical Structure to Therapeutic Applications. Edited by C.G. Knight. Pp.497. ISBN 0-444-80320-3. (Elsevier/North-Holland Biomedical: 1981.) \$110.75, Dfl. 260.

THE behaviour of suspensions of phospholipids in water (liposomes) has been under study for many years now, originally as a model for the lipid phase of biological membranes but more recently to investigate the possible therapeutic applications of liposomes in the delivery of drugs, enzymes and so on.

It was apparent by the late 1960s that the fact that lipids formed sealed vesicles when dispersed in water meant that they might be used to trap and protect unstable compounds from a harmful environment and so, for example, might enable the oral administration of compounds that would otherwise be inactivated in the gastrointestinal tract. One possible advantage of

liposomes for such a role is that since they are composed of "natural" compound, they might be expected to have little biological activity on their own. It has not turned out to be quite so simple. The problems, and attempts to overcome them, comprise the second part of this book.

Unfortunately, liposomes have antigenic properties, described in a chapter by Scherphof and co-workers. Patel and Ryman discuss the fate of liposomes when administered to animals, and Finkelstein and Weissmann describe attempts to target liposomes to specific tissues. An obvious problem raised in several chapters is that for liposomes to act as efficient drug delivery systems, drugs must not readily leak out of them. As discussed by Knight, this can in part be overcome by designing hydrophobic pro-drugs which are more likely to remain associated with the liposomes. In part also, it can be overcome by choosing components for the liposomes which will minimize their "fluidity". Nevertheless, the conclusion reached by Fildes is that the application of liposomes in the pharmaceutical industry is a long way off, because of a wide range of technical problems. At their most basic, liposomes would not be expected to last for very long in the average bathroom cabinet.

The overall impression of the second half of the book is then, as far as therapeutic applications are concerned, that liposomes are an answer awaiting a problem. With the first half of the book, we are on much more solid ground. This sets out to describe the physical properties of liposomes, with the aim of providing the information necessary for the proper design of liposomes for therapeutic applications. I imagine, however, that the book will be read mostly by those with an interest in biological membranes. It does provide an excellent, non-technical summary of our knowledge of the lipid bilayer, although the lack of discussion of the effects of proteins perhaps limits its general interest.

As is proper, the first chapter is a general introduction by Alex Bangham, who has done so much to generate the current interest in liposomes. A chapter by Eibl describes the synthetic routes to a wide range of phospholipids that he has developed based on reactions of phosphatidic acid dichlorides. We then have chapters on freeze-fracture studies, differential scanning calorimetry, X-ray, NMR and ESR. The latter chapter, as is appropriate for ESR, has an aura of mystery about it — the journal volume numbers given as $\zeta \ll \frac{1}{2}$, $\square \frac{1}{4}$, $\square \odot$ and even $\odot \frac{1}{3} \odot$ must be arcane cabalistic signs.

Finally, readers will have to decide for themselves whether or not Thudicum's description of lipids as "the centre, life and chemical soul of all bioplasm" (quoted by Bangham) is a little over the top. □

A.G. Lee is a Lecturer in the Department of Biochemistry at the University of Southampton.

A second edition of Sarnat and Netsky's *Evolution of the Nervous System* has been published by Oxford University Press, price £9.95. The first edition was reviewed in *Nature* 252, 428; 1974.

BOOKS RECEIVED

Mathematics

- CHEREMISINOFF, N.P. and CHEREMISINOFF, P.N. Unit Conversions and Formulas Manual. Pp.171. Flexi ISBN 0-250-40331-5. (Butterworth's: 1980.) £5.20.
- DAUBEN, J. W. (ed.). Mathematical Perspectives: Essays on Mathematics and Its Historical Development. Pp.272. ISBN 01-12-204050-3. (Academic: 1981.) \$34.
- FLOYD, A. (ed.). Developing Mathematical Thinking. Pp.294. Pbk ISBN 201-10237-4. (Addison-Wesley: 1981.) Pbk £6.95.
- GIBLIN, P.J. Graphs, Surfaces and Homology. An Introduction to Algebraic Topology. 2nd Edn. Pp.329. Pbk ISBN 0-412-23900-0. (Chapman & Hall: 1981.) Pbk £6.95.
- GORDON, A.D. Classification Methods for the Explanatory Analysis of Multivariate Data. Pp.193. ISBN 0-412-22850-5. (Chapman & Hall: 1981.) Np.
- GURTIN, M. E. An Introduction to Continuum Mechanics. Vol. 158, Mathematics in Science and Engineering. Pp.265. ISBN 0-12-309750-9. (Academic: 1981.) \$34.50.
- KUBERT, D.S. and LANG, S. Modular Units. Grundlehren der Mathematischen Wissenschaften 244. Pp.358. ISBN 3-540-90517-0. (Springer-Verlag: 1981.) DM 98, \$45.70.
- ONDAR, Kh. O. (ed.). The Correspondence Between A.A. Markov and A.A. Chuprov on the Theory of Probability and Mathematical Statistics. Pp.181. ISBN 3-540-90585-5. (Springer-Verlag: 1981.) DM 74, \$34.50.
- PRUGOVECKI, E. Quantum Mechanics in Hilbert Space, 2nd Edn. Pp.685. ISBN 0-12-566060-X. (Academic: 1981.) \$39.50.

Astronomy

- COCKS, V. D. Cosmos. Small Book, No. 3. Time, Space and Causality — and Other Matters. Pp.52. Pbk (no ISBN) (V. D. Cocks, Essex: 1981.) Pbk np.
- DUFFET-SMITH, P. Practical Astronomy with Your Calculator. 2nd Edn. Pp.188. Hbk ISBN 0-521-24059-X; pbk ISBN 0-521-28411-2. (Cambridge University Press: 1982.) Hbk £15; pbk £4.95.
- MOORE, P. (ed.). 1982 Yearbook of Astronomy. Pp.240. ISBN 0-393-01492-4. (W. W. Norton: 1982.) \$14.95.
- ORRAL, F. Q. (ed.). Solar Active Regions. A Monograph from Skylab Solar Workshop III. Pp.350. ISBN 0-87081-085-5. (Colorado Associated University Press: 1981.) \$17.50.

Physics

- ASHBY, M. F. *et al.* (eds). Dislocation Modelling of Physical Systems. Proceedings of the International Conference, Gainesville, Florida, U.S.A., June 1980. Pp.550. ISBN 0-08-026724-6. (Pergamon: 1981.) £31, \$75.
- BORISOVA, Z. U. Glassy Semiconductors. Pp.505. ISBN 0-306-40609-8. (Plenum: 1981.) \$69.50.
- BROGLIA, R. A. and WINTHER, A. Heavy Ion Reactions. Lecture Notes: Vol. I, Elastic and Inelastic Reactions. Frontiers in Physics, Vol. 52. Pp.291. ISBN 0-8053-1302-8. (Benjamin/Cummings Advanced Book Program, California: 1981.) \$29.50.
- CASTELLANI, C., DI CASTRO, C. and PELITI, L. (eds.). Disordered Systems and Localization. Lecture Notes in Physics, Vol. 149. Proceedings of the Conference held in Rome, May 1981. Pp.308. ISBN 3-540-11163-8. (Springer-Verlag: 1981.) DM38, \$17.70.
- CHEN, S.-H., CHU, B. and NOSSAL, R. (eds). Scattering Techniques Applied to Supramolecular and Nonequilibrium Systems. NATO Advanced Study Institutes Series, Vol. 73. Proceedings of a NATO Advanced Study Institute held August 1980 at Wellesley College, Wellesley, Massachusetts. Pp.928. ISBN 0-306-40828-7. (Plenum: 1981.) \$95.
- FABIAN, D. J., KLEINPOPPEN, H. and WATSON, L. M. (eds). Inner-Shell and X-Ray Physics of Atoms and Solids. Physics of Atoms and Molecules. Proceedings of the International Conference held August 1980 at Stirling, Scotland. Pp.950. ISBN 0-306-40819-8. (Plenum: 1981.) \$95.
- FONASH, S. J. Solar Cell Device Physics. Pp.332. ISBN 0-12-261980-3. (Academic: 1981.) \$35.
- GADIAN, D. G. Nuclear Magnetic Resonance and its Applications to Living Systems. Pp.197. ISBN 0-19-854627-0. (Oxford University Press: 1981.) £15.
- GORINGE, M. J. (ed.). Electron Microscopy and Analysis, 1981. Proceedings of the Institute of Physics Conference held at the University of Cambridge, September 1981 (EMAG 81). Conference Series No. 61. Pp.563. ISBN 0-85498-152-7. (The Institute of Physics, Bristol: 1982.) \$70 (US), £32 (elsewhere).
- GULDBERG, J. (ed.). Neutron-Transmutation-Doped Silicon. Proceedings of the 3rd International Conference held August 1980 in Copenhagen, Denmark. Pp.505. ISBN 0-306-04738-8. (Plenum: 1981.) \$59.50.
- GUSTAFSON, K. E. and REINHARDT (eds). Quantum Mechanics in Mathematics, Chemistry, and Physics. Proceedings of a Special Session organized as part of the 774th Meeting of the American Mathematical Society, held March 1980 in Boulder, Colorado. Pp.506. ISBN 0-306-40737-X. (Plenum: 1981.) \$59.50.
- HARTKÄMPER, A. and SCHMIDT, M.-J. (eds). Structure and Approximation in Physical Theories. Proceedings of a Colloquium held at Osnabrück, FRG, June 1980. Pp.255. ISBN 0-306-40882-1. (Plenum: 1981.) \$39.50.
- HEINZ, K. *et al.* (eds). Structural Studies of Surfaces. Springer Tracts in Modern Physics, Vol. 91. Pp.180. ISBN 3-540-10964-1. (Springer-Verlag: 1982.) DM65, \$30.30.
- KEESE, A. M., CONNOLLY, T. F. and BATTLE, G. C. Jr. (eds). Crystal Growth Bibliography Supplement. Solid State Physics Literature Guides, Vol. 11. Pp.262. ISBN 0-306-68331-8. (IFI/Plenum: 1981.) \$75.

Chemistry

- BOCKRIS, J. O'M. *et al.* (eds). Electrochemical Materials Science. Comprehensive Treatise of Electrochemistry, Vol.4. Pp.563. ISBN 0-306-40614-4. (Plenum: 1981.) \$55.
- CHAKRABARTI, C.L. (ed.) Progress in Analytical Atomic Spectroscopy, Vol.3. Pp.390. ISBN 0-08-029081-7. (Pergamon: 1981.) £45, \$105.
- MAMANTOV, G. and BRAUNSTEIN, J. (eds). Advances in Molten Salt Chemistry, Vol.4. Pp.444. ISBN 0-306-40833-3. (Plenum: 1981.) \$53.
- MILLER, J.N. (ed.). Standards in Fluorescence Spectrometry. Pp.115. ISBN 0-412-22500-X. (Chapman & Hall: 1981.) £8.50.
- O'KEEFE, M. and NAVROTSKY, A. (eds). Structure and Bonding in Crystals, Vol.1. Pp.327. ISBN 0-12-525101-7. (Academic: 1981.) \$48.
- PERRIN, D.D., DEMPSEY, B. and SERJEANT, E.P. *pKa* Prediction for Organic Acids and Bases. Pp.146. ISBN 0-412-22190-X. (Chapman & Hall: 1981.) £10.95.

Biological Sciences

- BEADLE, L.C. The Inland Waters of Tropical Africa. An Introduction to Tropical Limnology. 2nd Edn. Pp.475. ISBN 0-582-46341-6. (Longman: 1981.) £22.50.
- BELL, W.J. and ADIYODI, K.G. (eds). The American Cockroach. Pp.529. ISBN 0-412-16140-0. (Chapman & Hall: 1981.) £35.
- BERG, G.G. and MAILLIE, H.D. (eds). Measurement of Risks. Proceedings of the 13th Rochester International Conference on Environmental Toxicity, held June 1980, at the University of Rochester, New York. Conference CONF-800601. Pp.550. ISBN 0-306-40818-X. (Plenum: 1981.) \$59.50.
- BERRA, T.M. An Atlas of Distribution of the Freshwater Fish Families of the World. Pp.197. Hbk ISBN 0-8032-1411-1; pbk ISBN 8032-6059-8. (University of Nebraska Press: 1981.) Np.
- BOURNE, G.H., DANIELLI, J.F. and JEON, K.W. (eds). International Review of Cytology, Vol.73. Pp.365. ISBN 0-12-364473-9. (Academic: 1981.) \$45.
- BOWEN, I.D. and LOCKSHIN, R.A. (eds). Cell Death in Biology and Pathology. Pp.493. ISBN 0-412-16010-2. (Chapman & Hall: 1981.) £30.
- CIBA FOUNDATION SYMPOSIUM 86. The Fetus and Independent Life. Pp.372. ISBN 0-272-79650-6. (Pitman Books: 1981.) Np.
- DURIN, B. Insects, etc: An Anthology of Arthropods Featuring a Bounty of Beetles. Pp.108. ISBN 0-933920-25-3. (Hudson Hills Press, New York: 1981.) \$50.
- EAGLE, D.J. and CAVERLY, D.J. (eds). Diagnosis of Herbicide Damage to Crops. Reference Book 221. Pp.70. Pbk ISBN 0-11-240872-9. (HMSO: 1981.) Pbk £7.75.
- EATON, R.L. The Cheetah. Nature's Fastest Racer. Pp.128. ISBN 0-396-07994-6. (Dodd Mead, New York: 1981.) \$8.95.
- FARID, N.R. (ed.). HLA in Endocrine and Metabolic Disorders. Pp.357. ISBN 0-12-247780-4. (Academic: 1981.) \$39.
- FELTWELL, J. Large White Butterfly. The Biology, Biochemistry and Physiology of *Pieris brassicae* (Linnaeus). Series Entomologica, Vol. 18. Pp.542. ISBN 90-6193-128-8. (W. Junk, The Hague: 1981.) Dfl.225, \$98.
- FERMI, G. and PERUYZ, M.F. Atlas of Molecular Structures in Biology. 2, Haemoglobin and Myoglobin. Pp.104. ISBN 0-19-854706-4. (Oxford University Press: 1981.) £20.
- FIECHTER, A. (ed.). Advances in Biochemical Engineering. Vol.20, Bioenergy. Pp.209. ISBN 3-540-11018-6. (Springer-Verlag: 1981.) DM82, \$38.20.
- GARCIA, R.V. and ESCUDERO, J.C. Drought and Man. The 1972 Case History. Report on an IFIAS Project. Vol.2, The Constant Catastrophe: Malnutrition, Famines and Drought. Prepared at The Graduate Institute for International Studies, Geneva, 1976-1979. Pp.204. ISBN 0-08-025824-7. (Pergamon: 1981.) £20, \$45.
- GILBERT, L.I. and FRIEDEN, E. (eds). Metamorphosis. A Problem in Developmental Biology. 2nd Edn. Pp.578. ISBN 0-306-40692-2. (Plenum: 1981.) \$39.50.
- GILMAN, S., BLOEDEL, J.R. and LECHTENBERG, R. Disorders of the Cerebellum. Contemporary Neurology Series, Vol.21. Pp.415. ISBN 0-8036-4150-8. (F.A. Davis, Philadelphia: 1981.) \$40 (USA) \$48 (Canada).
- GLASS, R.E. Gene Function. *E.coli* and Its Heritable Elements. Pp.487. Hbk ISBN 0-7099-0081-3; pbk ISBN 0-7099-0082-1. (Croom Helm, London: 1981.) Hbk £19.95, pbk £9.95.
- GLEICHER, N. (ed.). Reproductive Immunology. Progress in Clinical and Biological Research, Vol.70. Including papers from the 1st S.B. Gusberg Seminar on Reproductive Immunology held at The Mount Sinai Medical Centre, New York, June 1980. Pp.510. ISBN 0-8451-0070-X. (Alan R. Liss, New York.) £58.40, DM257.
- GORROD, J.W. (ed.). Testing for Toxicity. Pp.381. ISBN 0-85066-218-4. (Taylor & Francis, London: 1981.) £18.
- GREGER, R., LANG, F. and SILBERNAGL, S. (eds). Renal Transport of Organic Substances. Pp.314. ISBN 3-540-10904-8. (Springer-Verlag: 1981.) DM74, \$34.50.
- GRIFFIN, D.H. Fungal Physiology. Pp.383. 0-471-05748-7. (Wiley: 1981.) £24.25, \$43.50.
- GRIMSON, W.E.L. From Images to Surfaces. A Computational Study of the Human Early Visual System. Pp.274. ISBN 0-262-07083-9. (MIT Press: 1981.) \$25.
- HADDEN, J.W. and STEWART, W.E. II. (eds). The Lymphokines Biochemistry and Biological Activity: Contemporary Immunology. Pp.437. ISBN 0-89603-012-1. (Humana Press, Clifton, New Jersey: 1981.) \$59.50. (US) \$69.50 (elsewhere).

HAMES, B.D. and RICKWOOD, D. (eds). *Gel Electrophoresis of Proteins: A Practical Approach*. Pp.290. Pbk ISBN 0-904147-22-3. (Information Retrieval, Oxford: 1981.) Pbk £7.50, \$18.

O'BRIEN, T.P. and McCULLY, M.E. *The Study of Plant Structure Principles and Selected Methods*. Pp.360. Pbk ISBN 0-9594174-0-0. (Blackwell Scientific: 1981.) Hbk £20; pbk £12.

PEPEU, G. and LADINSKY, H. (eds). *Cholinergic Mechanisms. Phylogenetic Aspects, Central and Peripheral Synapses, and Clinical Significance. Advances in Behavioral Biology, Vol.25. Proceedings of an International Symposium held March 1980 in Florence, Italy*. Pp.989. ISBN 0-306-40810-4. (Plenum: 1981.) \$95.

PISTORIUS, A. *The Country Journal Book of Birding and Bird Attraction*. Pp.274. ISBN 0-393-01493-2. (W.W. Norton: 1981.) \$15.95.

POLLACK, R. (ed.). *Readings in Mammalian Cell Culture*. 2nd Edn. Pp.720. Pbk ISBN 0-87969-137-9. (Cold Spring Harbor Laboratory, New York: 1981.) Pbk \$26 (US) \$31.20 (elsewhere).

PORTER, J.W. and SPURGEON, S.L. (eds). *Biosynthesis of Isoprenoid Compounds, Vol.1*. Pp.558. ISBN 0-471-04807-0. (Wiley: 1981.) £44, \$130.

PREBBLE, J.N. *Mitochondria, Chloroplasts and Bacterial Membranes*. Pp.378. ISBN 0-582-44133-1. (Longman: 1981.) £12.

RECHCIGL, M. (ed.). *Physiology of Growth and Nutrition. Comparative Animal Nutrition, Vol.4*. Pp.344. ISBN 3-8055-1199-X. (Karger, Basel: 1981.) SwFr.249, DM 299, \$149.25.

ROSE, D.A. and CHARLES-EDWARDS, D.A. (eds). *Mathematics and Plant Physiology. Experimental Botany. Proceedings of the 10th Annual Meeting of the Crop Science Model Builders' Group of the ARC, Glasshouse Crops Research Institute, March 1980*. Pp.320. ISBN 0-12-596880-9. (Academic: 1981.) £28.60, \$69.

ROUND, F.E. *The Ecology of Algae*. Pp.653. ISBN 0-521-22583-3. (Cambridge University Press: 1981.) £60.

SATO, V.L. and GEFTER, M.L. (eds). *Cellular Immunology: Selected Readings and Critical Commentary*. Pp.647. Pbk ISBN 0-201-10434-2. (Addison-Wesley: 1981.) Pbk \$21.50.

SCHACTER, A.J. (ed.). *Drug Dependence and Alcoholism. Biomedical Issues, Vol.1*. Pp.1342. ISBN 0-306-40323-4. (Plenum: 1981.) \$39.75.

SCOTT, W.N. and GOODMAN, D.B.P. *Hormonal Regulation of Epithelial Transport of Ions and Water. Annals of the New York Academy of Sciences, Vol.372*. Pp.660. Hbk ISBN 0-89766-133-8; pbk ISBN 0-89766-134-6. (The New York Academy of Sciences: 1981.) Hbk np; pbk \$140.

SIGEL, H. (ed.). *Metal Ions in Biological Systems. Vol.12, Properties of Copper*. Pp.384. ISBN 0-8247-1429-6. (Marcel Dekker: 1981.) SwFr. 165.

SINGLETON, P. and SAINSBURY, D. *Introduction to Bacteria for Students in the Biological Sciences*. Pp.167. Hbk ISBN 0-471-10034-X; pbk ISBN 0-471-10035-8. (Wiley: 1981.) Hbk £11, \$31; pbk £4.95, \$13.95.

STOKES, I.A.F. (ed.). *Mechanical Factors and the Skeleton*. Pp.224. ISBN 0-86196-006-8. (John Libbey, London: 1981.) £14.95.

UTTAL, W.R. *A Taxonomy of Visual Processes*. Pp.1097. ISBN 0-89859-075-2. (Lawrence Erlbaum, London: 1981.) £35.

VAN MAERCKE, Y.M.F., VAN MOER, E.M.J. and PELCKMANS, P.A.R. (eds). *Stomach Diseases Current Status. Proceedings of the 13th International Congress, Antwerp, May 1981*. Pp.434. ISBN 90-219-0500-0. (Excerpta Medica, Amsterdam: 1981.) \$76.50, Dfl. 180.

WATSON, L. *Sea Guide to Whales of the World: A Complete Guide to the World's Living Whales, Dolphins and Porpoises*. Pp.302. ISBN 0-09-146600-8. (Hutchinson, London: 1981.) £12.95.

WAXMAN, S.G. and RITCHIE, J.M. (eds). *Demyelinating Diseases. Basic and Clinical Electrophysiology. Advances in Neurology, Vol.31*. Pp.544. ISBN 0-89004-622-5. (Raven: 1981.) \$60.

WRONG, O.M., EDMONDS, C.J. and CHADWICK, V.S. *The Large Intestine: Its Role in Mammalian Nutrition and Homeostasis*. Pp. 217. ISBN 0-85200-351-4. (MTP, Lancaster, UK: 1981.) £16.95.

Applied Biological Sciences

BELFIORE, F. *Reviewed Contents of Major Diabetes Congresses. Frontiers in Diabetes, Vol.1*. Pp.162. ISBN 3-8055-3414-0. (Karger, Basel: 1981.) SwFr. 89, DM 107, \$53.50.

BERCHTOLD, P., SIMS, E.A.H. and BRANDEAU, K. (eds). *Obesity and Hypertension. International Journal of Obesity (1981), Suppl. 1, 1-188. Proceedings of a Satellite Symposium to the 3rd International Congress, held in Florence, October 1980*. Pp.192. (no ISBN) (John Libbey, London: 1981.) £14.95.

CHIRIGOS, M.A., MITCHELL, M., MASTRANGELO, M.J. and KRIM, M. (eds). *Mediation of Cellular Immunity in Cancer by Immune Modifiers. Progress in Cancer Research and Therapy, Vol.19*. Pp.287. ISBN 0-89004-628-y. (Raven: 1981.) \$31.50.

COLLEE, J.G. *Applied Medical Microbiology. Basic Microbiology, Vol.3. 2nd Edn*. Pp.150. Pbk ISBN 0-632-00853-9. (Blackwell Scientific: 1981.) Pbk £5.50.

FRIEDMAN, L.M., FURBERG, C.D. and DEMETS, D.L. *Fundamentals of Clinical Trials*. Pp.225. ISBN 0-88416-296-6. (John Wright, Bristol: 1981.) £13.

GARATTINI, S., GOLDIN, A., HAWKING, F. and KOPIN, I.J. (eds). *Advances in Pharmacology and Chemotherapy, Vol.18 — 1981*. Pp.357. ISBN 0-12-032918-2. (Academic: 1981.) \$45.

GOLDSTEIN, A.L. and CHIRIGOS, M.A. (eds). *Lymphokines and Thymic Hormones: Their Potential Utilization in Cancer Therapeutics. Progress in Cancer Research and Therapy, Vol.20*. Pp.342. ISBN 0-89004-697-2. (Raven: 1981.) \$38.

JUGE, O. and DONATH, A. (eds). *Neuronuclear Medicine. Progress in Nuclear Medicine, Vol.7*. Pp.176. ISBN 3-8055-2319-X. (Karger, Basel: 1981.) SwFr. 155, DM 186, \$93.

MESSIHA, F.S. and TYNER, G.S. (eds). *Endocrinological Aspects of Alcoholism. Progress in Biochemical Pharmacology, Vol.18. 4th Annual Conference, El Paso, Texas, February 1980*. Pp.232. ISBN 3-8055-2689-X. (Karger, Basel: 1981.) SwFr. 155, DM 186, \$93.

OXENDER, D. *et al.* (eds). *Membrane Transport and Neuroreceptors. Progress in Clinical and Biological Research, Vol.63. Proceedings of the ICN-UCLA Symposium held at Keystone, Colorado, February 1980*. Pp.470. ISBN 0-8451-0063-7. (Alan R. Liss, New York: 1981.) £46.40, DM 204.

PAIGE, D.M. and BAYLESS, T.M. (eds). *Lactose Digestion. Clinical and Nutritional Implications*. Pp.280. ISBN 0-8018-2647-0. (The Johns Hopkins University Press: 1981.) £15.75, \$28.10.

SANDLER, M. and USDIN, E. (eds). *Phenosulfotransferase in Mental Health Research*. Pp.225. ISBN 0-333-320166. (Macmillan Press, London & New York: 1981.) £22.

TURNER, M. and THOMAS, B. (eds). *Nutrition and Diabetes. Proceedings of a Scientific Meeting held at the University of Kent, Canterbury, July 1980*. Pp.112. ISBN 0-86196-008-4. (John Libbey, London: 1981.) £9.

USDIN, E. *et al.* (eds). *Clinical Pharmacology in Psychiatry: Neuroleptic and Antidepressant Research*. Pp.409. ISBN 0-333-31088-8. (Macmillan Press, London & New York: 1981.) £40.

WARDELL, W.M. and VELGO, G. (eds). *Drug Development, Regulatory Assessment, and Postmarketing Surveillance. Proceeding of a NATO Advanced Study Institute held October 1980 in Erice, Sicily*. Pp.356. ISBN 0-306-40822-8. (Plenum: 1981.) \$45.

General

ASHBY, E. and ANDERSON, M. *The Politics of Clean Air*. Pp.178. ISBN 0-19-858330-3. (Oxford University Press: 1981.) £15.

AUBURN, F.M. *Antarctic Law and Politics*. Pp.361. ISBN 0-905838-39-4. (C. Hurst, London: 1982.) £17.50.

BENHAM, H. *Man's Struggle for Food*. Pp.496. Hbk ISBN 0-8191-1518-5; pbk ISBN 0-8191-1519-3. (University Press of America, Lanham, Maryland: 1981.) Hbk \$26.50; pbk \$15.50.

BILLINGHAM, J. (ed.). *Life in the Universe*. Pp.461. Hbk ISBN 0-262-02155-2; pbk ISBN 0-262-52062-1. (MIT Press: 1982.) Hbk \$20; pbk \$12.50.

BOURNE, G.H. (ed.). *Physiology and Social Nutrition and Nutritional Education. World Review of Nutrition and Dietetics, Vol. 38*. Pp.230. ISBN 3-8055-3048-X. (Karger, Basel: 1981.) SwFr. 163, DM 195, \$97.75.

BROWN, A. *et al.* (eds). *The Cambridge Encyclopedia of Russia and the Soviet Union*. Pp.492. ISBN 0-521-23169-8. (Cambridge University Press: 1982.) £18.50.

BROWN, L.R. *Building a Sustainable Society*. Pp.433. ISBN 0-393-01482-7. (W.W. Norton: 1982.) £9.95, \$18.50.

CHANGEUX, J.-P. *et al.* *The Harvey Lectures. Delivered under the auspices of The Harvey Society of New York 1979-1980, Series 75*. Pp.430. ISBN 0-12-312075-6. (Academic: 1981.) \$48.50.

CHICKEN, J.C. *Nuclear Power Hazard Control Policy*. Pp.280. Hbk ISBN 0-08-023254-X; pbk ISBN 0-08-023255-8. (Pergamon: 1982.) Hbk £20, \$40; bpk £7.50, \$15.

COLLINS, H.M. and PINCH, T.J. *Frames of Meaning: The Social Construction of Extraordinary Science*. Pp.210. ISBN 0-7100-9011-0. (Routledge & Kegan Paul: 1982.) £12.50.

EWING, J. and KOSNIOWSKI, C. *Puzzle It Out. Cubes, Groups and Puzzles*. Pp.64. Pbk ISBN 0-521-28924-6. (Cambridge University Press: 1982.) Pbk £2.50.

GAL-OR, B. *Cosmology, Physics, and Philosophy*. Pp.522. ISBN 0-387-90581-2. (Springer-Verlag, New York: 1981.) Np.

GIBBS, L. *The Love Canal: My Story*. Pp.174. Hbk ISBN 0-87395-587-0; pbk ISBN 0-87395-588-9. (State University of New York Press: 1982.) Hbk \$11.95; pbk np.

GIBSON, R. *The Family Doctor: His Life and History*. Pp.214. ISBN 0-04-610017-2. (George Allen & Unwin: 1982.) £9.50.

GRAMMATICAKIS, P. *L'Idée d'indice de Conjugaison Relatif. Base d'un Essai d'une Cosmothéorie Pragmatique. Prolégomènes à la Métachimie et à l'Epostémologie*. Pp.70. Pbk ISBN 2-85206-128-7. (Librairie Scientifique et Technique Albert Blanchard, Paris: 1981.) Pbk np.

PONOMAREV, B.N. *Selected Speeches and Writings*. Pp.372. ISBN 0-08-023606-5. (Pergamon: 1981.) £19.75, \$47.50.

RAPHAEL, R. *Tree Talk. The People and Politics of Timber*. Pp.308. Flexi ISBN 0-933280-10-6. (Island Press, Covelo, California: 1981.) \$12 + sales tax.

RICHARDS, W.G. *et al.* *A Bibliography of Ab Initio Molecular Wave Functions: Supplement for 1978-80*. Pp.463. Flexi ISBN 0-19-1855367-6. (Clarendon Press/Oxford University Press: 1981.) £35.

STEWART, D.C. *Handling Radioactivity. A Practical Approach for Scientists and Engineers*. Pp.282. ISBN 0-471-04557-8. (Wiley: 1981.) £25.90, \$46.55.

THIRGOOD, J.V. *Man and the Mediterranean Forest. A History of Resource Depletion*. Pp.180. ISBN 0-12-687250-3. (Academic: 1981.) £12, \$29.

TIBBS, H. *The Future of Light*. Pp.103. Pbk ISBN 0-7224-0196-5. (Watkins, Dulverton, Somerset: 1981.) Pbk £2.95.

TONGREN, S. *What's for Lunch. Animal Feeding at the Zoo*. Pp.128. Hbk ISBN 0-939456-00-1; pbk ISBN 0-939456-01-X. (GMG Publishing, New York: 1981.) Hbk \$12.95; pbk \$8.95.

WHEELER, A. and PRICE, J.H. (eds). *History in the Service of Systematics. Papers from the Conference to celebrate the Centenary of the British Museum (Natural History) April 1981*. Pp.164. Pbk ISBN 0-90-184305-9. (British Museum (Natural History), London: 1981.) Pbk £10, \$23.

WILLIAMS, R. (ed.). *Contact: Human Communication and its History*. Pp.272. ISBN 0-500-01239-3. (Thames & Hudson: 1981.) £9.95.

WISPELAERE, C. De. (ed.). *Air Pollution Modeling and its Application, I. NATO — Challenges of Modern Society. Proceedings of the 11th International Meeting held November 1980, Amsterdam, The Netherlands*. Pp.747. ISBN 0-306-40820-1. (Plenum: 1981.) \$75.

WOLFF, M., NORTH, J.S. and DEERING, D. *The Waterloo Directory of Victorian Periodicals 1824-1900, Phase I*. Pp.1187. ISBN 0-08-026079-9. (Pergamon: 1981.) £88.95, \$200.